STUDIES in MATHEMATICAL STATISTICS

THEORY and APPLICATIONS

Edited by K. Sarkadi and I. Vincze

Akadémiai Kiadó, Budapest

STUDIES IN MATHEMATICAL STATISTICS

THEORY AND APPLICATIONS

Edited by

K. SARKADI and I. VINCZE

The theory of mathematical statistics has shown general progress during the last decades. In Hungary, where research in this line has valuable traditions, the Mathematical Institute of the Academy of Sciences set the aim of extending the application of mathematical statistics. In 1964 an international colloquium, organized by the János Bolyai Mathematical Society and the Scientific Society of Mechanical Engineers, attracted some 60 experts from 9 foreign countries and more than 200 Hungarian participants. The topics discussed covered several theoretical issues, such as the theory of probability, the problems of statistical methods of quality control and the question of life tests. The volume is composed partly of the lectures delivered at the colloquium, partly of the reports of Hungarian experts on their latest results.

AKADÉMIAI KIADÓ

PUBLISHING HOUSE
OF THE HUNGARIAN ACADEMY
OF SCIENCES

BUDAPEST

STUDIES IN MATHEMATICAL STATISTICS

THEORY AND APPLICATIONS

STUDIES IN MATHEMATICAL STATISTICS

THEORY AND APPLICATIONS

EDITED

BY

K. SARKADI AND I. VINCZE

AKADÉMIAI KIADÓ

PUBLISHING HOUSE OF THE HUNGARIAN ACADEMY OF SCIENCES

BUDAPEST 1968

© AKADÉMIAI KIADÓ, BUDAPEST 1968

PRINTED IN HUNGARY

CONTENTS

Preface 7

Voraussetzungen zur erfolgreichen Anwendung der statistischen Qualitätskontrolle
S. Benussi 9

Analysis of states of operation of two interdependent parallel systems
W. Bojarski, K. Wisyniewski and E. Fidelis 15

Statistische Lebensdauerkontrolle unter Voraussetzung einer Mischung von Weibull-Verteilungen
O. Bunke 21

On the number of intersections between two empirical distribution functions
E. Csáki 27

Die Anwendung der C-Karte zur Kontrolle von Ziehteilen aus Feinblechen
B. Czyżewski 33

Die Steuerung der Qualität von Normteilen mit mathematisch-statistischen Methoden
H. Finsterbusch 41

Locally most powerful rank tests of independence
J. Hájek 45

Some new results in the theory of rank tests
J. Hájek 53

Allgemeine Grenzwertsätze der Wahrscheinlichkeitsrechnung
H. Heyer 57

Statistische Stichproben von langgestreckten Materialen
J. Křepela und P. Mandl 75

Anwendungsmöglichkeiten der Simulationstechnik im Textilmaschinenbau und in der Textilindustrie
U. Liebscher 81

Asymptotically optimal tests of composite hypotheses for randomized experiments with non-controlled predictor variables
J. Neyman and E. L. Scott 89

Über nichtlineare Optimierung
J. Oderfeld 101

On the connectedness of random graphs
I. Palásti 105

Definition of most favourable characteristics of tricot fabrics produced on interlock machines
N. Rancu, L. Tövissi and I. Teodorescu 109

Ein Problem der linearen Regulierung bei einer nicthstationären Störung von Polynomform
Z. Režný 119

Information and statistics
A. Rényi 129

Some remarks on the estimation of fraction defective and reliability in the case of normal distribution
K. Sarkadi 133

Experimentelle Untersuchung von Einfach-, Doppel- und Mehrfach-stichprobeplänen des Mil. Std. 105
O. Schürz 139

Additive Prozesse mit Zufallssprüngen in den Absorptionswall und Lebensdauermodelle bei den Brucherscheinungen in starren Körpern welche diesen Prozessen entsprechen
J. Sedláček 149

On the mean number and size of opaque particles in transparent bodies
Z. Šidák 161

Die wirtschaftlichste Verteilung des Prüfaufwands auf die verschiedenen Stufen eines Modells mit mehreren Zufallskomponenten
K. Stange 169

Die Berechnung und praktische Anwendung von Doppel- und Mehrfachstichprobenplänen für Variablenprüfung
H. Trampel 183

On the OC-function of the acceptance plan designed for a Pólya process of defective items
G. Tusnády 193

On the power of the Kolmogorov–Smirnov two-sample test and related non-parametric tests
I. Vincze 201

PREFACE

The theory of mathematical statistics has made considerable progress in the last decades. Hungary has traditions in this branch of science: Charles Jordan (1870–1959) did his pioneering work in Hungary; his books entitled *Mathematical Statistics* appeared in Paris and in Budapest, in the late twenties. Jordan also dealt with the practical applications of statistics, particularly in geophysics and meteorology.

The Mathematical Institute of the Hungarian Academy of Sciences has been engaged in research on mathematical statistics for about 15 years, developing the field and extending its application. As a result of this work, statistical methods of quality control have been introduced in a number of manufacturing processes in Hungary, and mathematical statistics find increasing use in research institutes and industrial laboratories to design experiments and evaluate the data. Several research institutes and other industrial organizations employ mathematical statisticians on their staff.

In 1964 the time was come to hold a Symposium on mathematical statistics in Hungary with international participation. This Symposium was jointly organized by the *János Bolyai Mathematical Society* and the *Scientific Society of Mechanical Engineers*. About 60 representatives of nine foreign countries were present at the Symposium; the number of Hungarian participants exceeded, together with the engineers and other people interested in the applications of mathematical statistics, two hundred persons.

The programme of the Symposium included several theoretical questions of mathematical statistics, the theory of reliability, the problems of statistical methods of quality control, and the problem of life tests.

This volume contains studies based on the most important lectures presented at the Symposium. The book has been complemented with several papers by Hungarian authors reporting recent results in researches on mathematical statistics.

K. Sarkadi *I. Vincze*

VORAUSSETZUNGEN ZUR ERFOLGREICHEN ANWENDUNG DER STATISTISCHEN QUALITÄTSKONTROLLE

von

S. BENUSSI

Kugellagerfabrik
Budapest, Ungarn

Die große Entwicklung der Produktion der ungarischen Wälzlagerindustrie hat die Voraussetzungen für die Fertigung von Wälzlagerbestandteilen in großen Serien geschaffen. Die fertigungstechnischen Gegebenheiten boten den Anlaß und die Möglichkeit zur Einleitung der modernsten Kontrollmethoden, des Systems der statistischen Qualitätskontrolle.

Auf Grund der aus der ungarischen und internationalen Literatur erworbenen Erfahrungen wurden bereits im Jahre 1958 die ersten Schritte zur Vorbereitung der Einleitung der statistischen Qualitätskontrolle unternommen. Während der seither verflossenen Zeit wurden verschiedene Arten der statistischen Qualitätskontrolle eingeführt: die statistische Produktionskontrolle und die Endkontrolle durch Stichprobenentnahme. Die im Zusammenhang mit der Einführung dieser Methoden erworbenen Erfahrungen bestätigten, daß die erfolgreiche Anwendung der statistischen Qualitätskontrolle die Erfüllung mehrerer Bedingungen voraussetzt. Das Einhalten dieser Bedingungen gewährleistet die mit der Anwendung der statistischen Qualitätskontrolle verbundenen außerordentlichen Erfolge. Unseres Erachtens können die Vorussetzungen der erfolgreichen Anwendung der statistischen Qualitätskontrolle in folgenden Punkten zusammengefaßt werden:

1. Technische Vorbereitung
2. Sicherung des materiellen Anreizes
3. Ermittlung und Beseitigung der Verlustquellen.

1. TECHNISCHE VORBEREITUNG

Eine der wichtigsten Aufgaben bei der Einführung der statistischen Qualitätskontrolle ist die Auswahl eines entsprechenden Gebietes.Dazu wurden in der Wälzlagerfabrik Diósd auf Grund der Fertigung und der technologischen Bedingungen die Dreherei und innerhalb dieser die Mehrspindelautomaten sowie die halbautomatischen Maschinen ausgewählt. Den zu verrichtenden Aufgaben entsprechend wurde entschieden, von den Methoden der statistischen Qualitätskontrolle sich in erster Linie mit der statistischen Produktionskontrolle zu befassen und erst später auf Grund der erworbenen Erfahrungen die Einleitung der Endkontrolle mittels Stichprobenentnahme vorzubereiten. Um dem gestellten Ziel näherzukommen wurde eine Studie erarbeitet, die sowohl die theoretischen als auch die

praktischen Fragen der anwendbaren statistischen Qualitätskontrolle erfaßte. Bei der Ausarbeitung dieser Studie wurden die Gegebenheiten der Fertigung sowie die fachliche Bereitschaft und Schulbildung der Werktätigen zugrunde gelegt und unter Berücksichtigung dieser Faktoren die theoretischen Begriffe einfach, die statistischen Kennwerte hingegen in der Weise festgelegt, daß letztere ohne besondere Berechnungsgänge durch einfache Abzählung ermittelt werden können.

Zwecks Erleichterung der Einleitung der statistischen Produktionskontrolle wurden nur jene wichtigsten Einflußgrößen der bearbeiteten Teile durch Beobachtung der statistischen Kennwerte kontrolliert, die in größerem Maße die nächstfolgenden Bearbeitungsgänge beeinflußten. Diese Einflußgrößen wurden mit den Kennwerten »Median« und »Spannweite« untersucht. Weitere weniger wichtige Einflußgrößen wurden lediglich qualifiziert. Die Qualifizierung erfolgt in den im unteren Teil des Kontrolldiagrammes angeordneten Quadraten unter der Anwendung einheitlich ausgewählter Bezeichnungen.

Dem Obengenannten gemäß wurde schließlich das zweckdienlichste Kontrolldiagramm angefertigt. In diesem wurden der Wert »Median« durch Kontrollgrenzen, der Wert »Spannweite« hingegen einfach durch eine absolute Zahl bezeichnet. Die Evidenzhaltung der Spannweite mit Hilfe einer absoluten Zahl ist nämlich einfacher und erleichtert die Handhabung des Kontrolldiagrammes.* Nun wurde der Lehrstoff der statistischen Qualitätskontrolle zusammengestellt und in einem 60stündigen Lehrgang der zur Durchführung der praktischen Aufgaben ausgewählten, aus 4 Personen bestehenden Gruppe vorgetragen. Der Stoff des Lehrganges umfaßte folgende Fragen:

1 Begriff der Häufigkeit und das Häufigkeitsdiagramm
2. Statistische Kennwerte (Durchschnitt, Median, Spannweite, Streuung)
3. Datenermittlung
4. Toleranz und Prüfung der Bearbeitungsgenauigkeit
5. Kontrollgrenze für Median und Spannweite
6. Anfertigung und Handhabung des Kontrolldiagrammes
7. Kontrolle und Bewertung der Fertigung auf Grund des Kontrolldiagrammes
8. Ermittlung des Einstellniveaux von spanabhebenden Werkzeugmaschinen
9. Endkontrolle durch Stichprobenentnahme.

Die in dieser Weise gebildete kleine Gruppe stellte nun den Kern der statistischen Qualitätskontrolle dar. Dieser Kern wurde ohne Personalerhöhung aus der technischen Kontrollabteilung ausgewählt. Einer der Ausgewählten war ein Techniker, die übrigen drei waren über mehrjährige Fachpraxis verfügende technische Kontrolleure. Der Techniker wurde später mit der Durchführung der theoretischen und administrativen Aufgaben der statistischen Qualitätskontrolle und mit der Leitung der Kontrollgruppe beauftragt.

* In diesem Zusammenhang verweisen wir auf: *Wissenschaftliche Zeitschrift der TH Magdeburg* **2**, 6 (1962).

Nach Absolvierung des Lehrganges setzten die 4 Kontrolleure die Aneignung der praktischen Kenntnisse der statistischen Produktionskontrolle in der Dreherei fort. Das Einüben dauerte ein halbes Jahr und wurde auf 5 Mehrspindelautomaten durchgeführt. Während dieser Zeit wurde für die übrigen Kontrolleure (laufende und Endkontrolle) ein 80stündiger Lehrgang abgehalten. An diesem Lehrgang wurden die technischen Kontrolleure auf Grund des vorangehend mitgeteilten Lehrstoffes zur Durchführung der Aufgaben der statistischen Qualitätskontrolle vorbereitet. Nach Beendigung des Lehrganges wurde mit Hilfe der mathematisch-statistischen Gruppe die statistische Produktionskontrolle in der Dreherei eingeleitet.

Durch Vermittlung der theoretischen und praktischen Kenntnisse konnte mit Hilfe der neuen Methode die Anzahl der kontrollierten Maschinen der Dreherei von 5 im Jahre 1959 auf 56 im Jahre 1962 erhöht werden. Die allmähliche, jedoch in einem langsamen Takt erfolgende Einleitung der statistischen Produktionskontrolle war durch die Schulung der Werktätige der Dreherei bedingt. Es war die Aufgabe der technischen Kontrolleure, die Werktätigen der Dreherei mit den Begriffen der statistischen Produktionskontrolle bekanntzumachen.

In dieser Weise wurde das Gebiet der statistischen Qualitätskontrolle stets erweitert. Die erworbenen Erfahrungen sowie die vorangegangenen Studien ermöglichten die Organisation der Endkontrolle durch Stichprobenentnahme. Bereits im Jahre 1960 wurden sämtliche im Rahmen der statistischen Produktionskontrolle geprüften Teile in der Endkontrolle mit dem Stichprobenverfahren qualifiziert. Zur Endkontrolle mittels Stichprobenentnahme wurde auf Grund der Ungarischen Norm Nr. MNOSZ 247–55 der zweifache Plantyp erwählt. Die zur Endkontrolle durch Stichprobenentnahme erforderliche Tabelle wurde durch die Ermittlung folgender Angaben festgelegt:

Übliche Losgrößen: untere Grenze 200 Stück
obere Grenze 3200 Stück

Genauigkeitsstufe: III

Fehlerprozentsatz (Nennwert): 0,5

Unter Verwendung dieser Angaben wurde Tabelle I zusammengestellt:

TABELLE I

Losgröße	Erste Probengruppe			In der zweiten Probengruppe zu entnehmende Stückzahl	In den beiden Probengruppen unzulässige fehlerhafte Stücke
	zu entnehmende Stückzahl	zulässige fehlerhafte Stücke	unzulässige fehlerhafte Stücke		
200–500	25	0	2	50	2
501–800	35	0	2	70	2
801–1200	50	0	3	100	3
1201–3200	75	1	3	150	3

Die allmähliche Einführung der 100%igen Endkontrolle und die Einleitung der Endkontrolle mittels Stichprobenentnahme ermöglichte eine weitgehende Herabsetzung der Zahl der technischen Kontrolleure. Die gemeinsame Anwendung der statistischen Produktionskontrolle und der Endkontrolle mittels Stichprobenentnahme erbrachte bereits im Jahre 1961 außergewöhnliche Erfolge. Die erreichte Wirtschaftlichkeit bestätigte den Nutzen und die Notwendigkeit der Weiterentwicklung der statistischen Qualitätskontrolle auf allen Gebieten.

2. SICHERUNG DES MATERIELLEN ANREIZES

Da anfangs sämtliche mit der statistischen Qualitätskontrolle verbundenen Aufgaben durch die Werktätigen der technischen Kontrollabteilung durchgeführt wurden, schritt die Einführung des neuen Systems nur langsam vorwärts. Die Aufklärung der Werktätigen und die Bekanntmachung der neuen Begriffe erforderte eine langwierige und zähe Arbeit. Die Werktätigen standen dem Kennenlernen der für sie vollkommen unbekannten statistischen Qualitätskontrolle sowie ihrer praktischen Anwendung anfänglich mit einigem Befremden gegenüber. Deshalb unternahm die Betriebsleitung die erforderlichen Schritte, um die in der Produktion tätigen Arbeiter für die Einleitung der statistischen Qualitätskontrollen zu interessieren. Es wurde der Grundsatz des materiellen Anreizes berücksichtigt, der Ausschußprozentsatz der ausschußgefährlichen Arbeitsgänge festgelegt und die Beurteilung der Arbeit unserer Werktätigen an die Erfüllung des Ausschußprozentsatzes gebunden. Arbeitsgruppen wurden für die einzelnen Arbeitsgänge mit Einstellern an ihrer Spitze gebildet. Die Arbeit dieser Gruppen wurde mit Hilfe der Endkontrolle durch die Methode der Stichprobenentnahme kontrolliert. Unterschreitet die Arbeitsgruppe den vorgeschriebenen Prozentsatz des Ausschusses (das vorgeschriebene Niveau), so erhalten die Werktätigen der Gruppe Prämien. Bereits 1961 wurden die Werktätigen, die bei ihrer Arbeit den vorgeschriebenen Qualitätsanforderungen genügten, regelmäßig prämiiert. Durch Einführung des Prämiensystems verdienen die besser arbeitenden Werktätigen des Betriebes um etwa 7–10% mehr als die übrigen Werktätigen.

Die einheitliche Zusammenfassung der Kräfte und die aktive Arbeit der Werktätigen leisteten dem Betrieb in der Einführung und Verbreitung der statistischen Qualitätskontrolle große Hilfe.

3. ERMITTLUNG UND BESEITIGUNG DER VERLUSTQUELLEN

Die Sicherung der Güte der einzelnen Teile erfordert sowohl von den Werktätigen in der Produktion als auch in anderen Abteilungen eine einheitliche Anschauung. Die Lösung der Aufgaben bei einheitlichem Standpunkt erleichtert die technische Anleitung und schafft die Grundlagen der persönlichen Verantwortung. Auf diesem Gebiet bietet die statistische Qualitätskontrolle einen besonderen Vorteil. Ein fachgemäß geführtes und ausgewertetes Kontrollkennbild gewährt sowohl für die technische Kontrolle als auch für die Betriebsleitung eine sichere Grundlage für das Ergreifen

von erforderlichen Maßnahmen. Die Untersuchung der Grundgenauigkeit der spanabhebenden Werkzeugmaschinen und die Beseitigung der im Wege der Fertigung liegenden Hindernisse bedeutet sowohl für die verfahrenstechnischen Organe als auch für sämtliche sonstigen technischen Stellen eine äußerst nützliche Hilfe. Eben deshalb werden die Kontrolldiagramme mit besonderer Rücksicht auf die ausschußgefährlichen Arbeitsgänge systematisch und monatlich ausgewertet. Die erworbenen Erfahrungen werden in Form eines analytischen Berichtes mit dem Oberingenieur des Betriebes mitgeteilt, der dann in der Lage ist, die erforderlichen Maßnahmen zu treffen. In dem üblichen analytischen Bericht wurde bereits auf die kurze Standzeit der Meißel, auf die Fehler der Spanvorrichtungen, auf die unzulängliche Grundgenauigkeit der spanabhebenden Werkzeugmaschinen und schließlich auf das unrichtig gewählte Einstellniveau der zur spanabhebenden Bearbeitung eingesetzten Maschinen hingewiesen.

Die durch die technische Kontrollabteilung erstellten analytischen Berichte ermöglichen die endgültige Ausschaltung der Verlustquellen auf organisatorischem Wege. In dieser Weise trägt die statistische Qualitätskontrolle zur Steigerung einer organisierten Fertigung und der Betriebssicherheit bei. Die rechtzeitige Behebung der Mängel in der Fertigung führt nicht nur zu einer Herabsetzung der anfallenden Ausschußmenge, sondern trägt auch zur Erhöhung des technischen Niveaus in großem Maße bei. Die durch die statistische Qualitätskontrolle erreichten Erfolge haben wir in Tabelle II zusammengefaßt.

TABELLE II

Jahr	Dreherei Produktionsplan %	Personalbestand %	Technische Kontrollabteilung		
			ohne statistische Kontrolle %	Personalabstand mit statistischer Kontrolle %	Ausschuß (Forint) %
1959	100	100	100	100	100
1960	114	115	110	110	83
1961	133	112	126	79	51
1962	152	119	142	63	32

Aus Tabelle II geht die Entwicklung der einzelnen Angaben, bezogen auf das Grundniveau des Jahres 1959, hervor. Aus der Untersuchung der Ergebnisse kann folgendes festgestellt werden:

1. Der mengenmäßig angegebene Produktionsplan der Dreherei ist im Jahre 1962 im Vergleich zum Grundniveau des Jahres 1959 um 52% angestiegen. Demgegenüber wurde die Belegschaft des Betriebes nur um 19% erhöht (siehe Spalte 2 und 3). Der geringere Zuwachs im Personalstand des Betriebes wurde durch den Einsatz moderner Maschinen und die Einführung der gleichzeitigen Zweimaschinen-Bedienung ermöglicht.

2. Die qualitative Erfüllung des erhöhten quantitativen Planes hätte bei der Anwendung des früheren Kontrollsystems eine Personalstanderhöhung

der Gütekontrolle um 42% erfordert. Da eine Personalstanderhöhung durch die Wirtschaftlichkeitskennwerte des Betriebes keineswegs begründet war, wurde die Einführung der die Herabsetzung des Personalstandes ermöglichenden statistischen Qualitätskontrolle zur zwangsläufigen Notwendigkeit.

3. Die gemeinsame Einführung der statistischen Produktionskontrolle sowie der Endkontrolle mittels Stichprobenentnahme ermöglichte eine allmähliche Herabsetzung des Personalstandes der technischen Kontrollableitung. Dies wird durch die im Jahre 1961 erreichte 21%ige und die im Jahre 1962 verwirklichte 37%ige Personalstandverminderung bestätigt (siehe Spalte 5).

4. Trotz der Erhöhung des Produktionsplanes und der Verringerung des Bestandes der technischen Kontrollabteilung trat in der Qualität der Erzeugnisse des Betriebes keine Verschlechterung ein. Die Entwicklung des Ausschußprozentsatzes zeigte ebenfalls einen vorteilhaften Verlauf. Als Folge der organisierten Arbeit ging der Ausschußprozentsatz im Jahr 1962 bereits um 68% zurück.

ANALYSIS OF STATES OF OPERATION OF TWO INTERDEPENDENT PARALLEL SYSTEMS

by

W. BOJARSKI, K. WISYNIEWSKI and E. FIDELIS

Mathematical Institute of the Polish Academy of Sciences
Warsaw, Poland

1. CHARACTERIZATION OF THE INTERDEPENDENCE OF THE SYSTEMS

Let us consider two systems of equipments, A and B. Systems A and B consist of N_A and N_B equipments connected in parallel, respectively. Equipments belonging to the same system are of the same kind, but those of different systems usually differ. System A, connected with system B by a single line, forms system P. Systems A and B are interdependent.

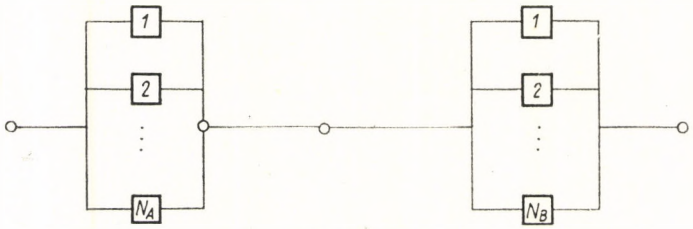


FIG. 1

Whenever one of the equipments belonging to system A fails, an equipment belonging to system B will be switched off, and inversely. In case of having failed equipments, belonging either to system A or to system B, they will be repaired. The equipments belonging to system B, switched off as a result of failures that have occured in system A, get into the reserve of system B. The equipments belonging to system A, switched off as a result of failures that have occured in system B, get into the reserve of system A. An equipment belonging to system A, being in reserve, is to be switched on if one of the operating equipments of system A has failed, and, at the same time, one equipment of system B having been under repair, is found to be repaired. Similarly, an equipment belonging to system B, being in reserve, is to be switched on if one of the operating equipments of system B has failed and, at the same time, one equipment of system A having been under repair, is found to be repaired.

Any state of the system A, denoted by $S_A(X_A, Y_A, Z_A)$ is characterized by the following three variables:

X_A = number of operating equipments

Y_A = number of equipments being under repair

Z_A = number of equipments being in reserve.

Similarly, any state of the system B, denoted by $S_B(X_B, Y_B, Z_B)$ is characterized by the following three variables:

X_B = number of operating equipments

Y_B = number of equipments being under repair

Z_B = number of equpiments being in reserve.

Any state of system P, denoted by $S_P(X_A, Y_A, Z_A; X_B, Y_B, Z_B)$ is characterized by the six variables: X_A, Y_A, Z_A, X_B, Y_B, Z_B, that must satisfy the following relations:

1. $X_A + Y_A + Z_A = N_A$
2. $X_B + Y_B + Z_B = N_B$
3. $X_A = X_B$
4. $Z_A Z_B = 0$.

These relations are derived from the conditions characterizing the interdependence of the systems A and B, being imposed on the cooperating work of the two systems.

The possible values of the random variables X_A, Y_A, Z_A and X_B, Y_B, Z_B are 0, 1, 2, . . ., N_A and 0, 1, 2 . . ., N_B, respectively.

By considering the interdependence of the system A and B it can be shown that the number of possible states of system P is given by

$$L = (N_A + 1)(N_B + 1).$$

All the states of system P are transient. In such formulation of the interdependence of the systems A and B it is assumed that the equipments belonging to system A are of the same kind, being derived from the same population of equipments. The equipments belonging to system B may have been derived from another population.

In order to simplify the cooperated work of systems A and B it is assumed that a failure of any equipment belonging to system A results in switching off one of the equipments belonging to system B, and inversely.

For instance, this scheme is effective for those cases where every equipment belonging to system A produces the same quantity of a certain product, while every equipment belonging to system B consumes just the same quantity of that product.

2. ASSUMPTIONS

If the probability of sojourning in a given state is known for any state of system P, conclusions can be drawn about the efficiency of the work of that system.

In order to evaluate those probabilities we make the following assumptions:

1. The failure rate of equipments belonging to system A, denoted by λ_A, is known. The (conditional) probability that one equipment out of X_A will fail within a time Δt, under the condition that all of them are operating at the moment t, is given by

$$X_A \lambda_A \Delta t + o(\Delta t) .$$

2. The intensity of repair of equipments belonging to system A, denoted by μ_A, is known. The (conditional) probability that one equipment out of Y_A ones will be repaired within a time Δt, under the condition that all of them are under repair at the moment t, is given by

$$Y_A \mu_A \Delta t + o(\Delta t) .$$

The constancy of the quantity μ_A is a consequence of the assumption that failures are of the same sort and that the number of repairing teams is large enough.

3. The failure rate of the equipments being in the reserve of system A, denoted by ν_A, is known. The (conditional) probability that one equipment out of Z_A will fail within a time Δt, under the condition that all of them are undamaged at the moment t, is given by

$$Z_A \nu_A \Delta t + o(\Delta t) .$$

Quite analogous assumptions are made on the equipments belonging to system B, under the condition that the intensities λ_B, μ_B, ν_B are known quantities.

3. EVALUATION OF THE PROBABILITY OF SOJOURNING IN A GIVEN STATE OF SYSTEM P

Any given state is attainable from two, three or, at most, four other states. This conclusion is drawn by an analysis of all possible transitions to a given state and by considering the conditions characterizing the interdependence of systems A and B.

Every state of system P is characterized by one of the three possible variants depending on the values of the variables Z_A and Z_B.

1. $Z_A = 0$, $Z_B \neq 0$,
2. $Z_A \neq 0$, $Z_B = 0$,
3. $Z_A = 0$, $Z_B = 0$.

The state denoted by the symbol

$$S_P^{(i)}(X_A, Y_A, 0 ; X_B, Y_B, Z_B) \qquad (\text{for } Z_A = 0 , Z_B \neq 0)$$

is only attainable from the states denoted by

$$S_P^{(j)}(X_A+1, Y_A-1, 0;\ X_B+1,\ Y_B,\ Z_B-1),$$
$$S_P^{(k)}(X_A-1, Y_A+1, 0;\ X_B-1,\ Y_B,\ Z_B+1),$$
$$S_P^{(l)}(X_A,\ Y_A,\ 0;\ X_B,\ Y_B+1,\ Z_B-1),$$
$$S_P^{(m)}(X_A,\ Y_A,\ 0;\ X_B,\ Y_B-1,\ Z_B+1).$$

Marking the states by i, j, k, l, m is corresponding to marking of probabilities of sojourning in a given state of system P with $P_i(t)$, $P_j(t)$, $P_k(t)$, $P_l(t)$, $P_m(t)$.

Let us form the probability $P_i(t+\Delta t)$ of sojourning in the state $S_P^{(i)}(X_A, Y_A, 0;\ X_B, Y_B, Z_B)$ of system P at the moment $t+\Delta t$.

By the limiting transition $\Delta t \to 0$ the following differential equation is obtained:

$$\frac{dP_i(t)}{dt} = -(X_A\lambda_A + Y_A\mu_A + X_B\lambda_B + Y_B\mu_B + Z_B\nu_B)P_i(t) +$$

$$+(X_A+1)\lambda_A P_j(t) + (Y_A+1)\mu_A P_k(t) + (Y_B+1)\mu_B P_l(t) + (Z_B+1)\nu_B P_m(t).$$

Analogous differential equations can be deduced for any of the L possible states of system P.

The problem consists in solving a system of L linear differential equations in which the unknown quantities are the probabilities of sojourning in any of the states of system P.

The value of the determinant of the characteristic equation for that system of L differential equations equals the product of elements of the principal diagonal. For that reason the roots of the characteristic equation can easily be determined. We adopted the initial condition that the probability of sojourning in state $S_P(N, 0, 0; N, 0, 0)$ of system P [where $N = \min(N_A, N_B)$] equals 1 at the moment $t = 0$. Thereof, the probability of sojourning in any other state of system P is equal to 0 at the moment $t = 0$. In this way the coefficients occuring in the equations relating to the probabilities, can be determined.

4. CONCLUSIONS

By summing up the obtained probabilities the following distributions can be calculated:

(a) the distribution of number X_A of working equipments belonging to system A, that is, at the same time, the distribution of number X_B of working equipments belonging to system B;

(b) the distribution of number Y_A of equipments belonging to system A and being under repair, and that of number Y_B of equipments belonging to system B and being under repair;

(c) the distribution of number Z_A of equipments being in the reserve of system A, and that of number Z_B of equipments being in the reserve of system B.

In the knowledge of these distributions an analysis may be carried out in concern with the work of systems of type P depending on the parameters λ_A, μ_A, ν_A, λ_B, μ_B, ν_B and on the numbers N_A and N_B.

Our present work is concerned with the investigation of more general systems, in which the failing of R equipments of system A need not be followed by switching off the same number of equipments of system B.

STATISTISCHE LEBENSDAUERKONTROLLE UNTER VORAUSSETZUNG EINER MISCHUNG VON WEIBULL-VERTEILUNGEN

von

O. BUNKE

Institut für Angewandte Mathematik und Mechanik
Deutsche Akademie der Wissenschaften zu Berlin
Berlin, Deutsche Demokratische Republik

1.

In Verallgemeinerung des einfachen Ansatzes einer Exponentialverteilung als Modell für die statistische Verteilung der Lebensdauer von elektronischen Elementen, insbesondere von Halbleitern und Elektronenröhren, ist von mehreren Autoren der Ansatz einer Weibull-Verteilung vorgeschlagen und untersucht worden [1, 2, 3]. In einer Reihe von Veröffentlichungen [4, 5, 6] haben Goode und Kao Methoden und Tabellen zur statistischen Kontrolle der mittleren Lebensdauer unter Voraussetzung einer Weibull-Verteilung bei bekanntem Gestaltsparameter β angegeben. Über den Gestaltsparameter werden oft nur ungenaue Erfahrungswerte vorliegen, wodurch Verfälschungen auftreten können [7, 11]. Im Abschnitt 4. werden wir ein Verfahren zur Lebensdauerkontrolle angeben, welches keine Kenntnis von β voraussetzt. In vielen Fällen, insbesondere bei einer Überlagerung verschiedener Ursachen für den Ausfall der Elemente und bei einer Vermischung von Elementen aus verschiedenen Posten, ist eine Mischung von Exponential- oder Weibull-Verteilungen als Modell für die Verteilung der Lebensdauer geeignet [1, 8]. Wir werden im folgenden für dieses Modell ein Verfahren zur statistischen Kontrolle der Lebensdauer angeben, bei dem die Anzahl z der Ausfälle von n Elementen während einer vorgegebenen Prüfzeit T zugrundegelegt wird.

2.

Eine Weibull-Verteilung mit dem Erwartungswert μ und dem Gestaltsparameter β hat die Wahrscheinlichkeitsdichte*

$$p(t \mid \mu,\beta) = \begin{cases} 0 & : t < 0 \\ \dfrac{\beta}{\alpha} t^{\beta-1} \exp\left[-\dfrac{t^\beta}{\alpha}\right] & : t \geq 0 \end{cases} \qquad \left(\alpha = \mu^\beta \left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^{-\beta}\right)$$

Wir wollen nun eine Mischung von drei Weibull-Verteilungen

$$\sum_{i=1}^{3} p_i p(t \mid \mu_i, \beta_i) \qquad [p_i \geq 0,\ \sum_{i=1}^{3} p_i = 1]$$

* Hierbei ist also $\mathbf{P}\{t \leq 0\} = 0$ vorausgesetzt.

als Verteilung der Lebensdauer ansetzen. Über die Erwartungswerte μ_i sei bekannt, daß sie in Intervallen $[\underline{\mu}_i, \bar{\mu}_i]$ liegen. Es sei bekannt, daß die Gestaltsparameter β_i im Intervall $[\underline{\beta}, \bar{\beta}]$ liegen. n Elemente, welche aus dem zu kontrollierenden Posten zufällig entnommen sind, werden während der Zeitdauer T geprüft. Ist die Anzahl z der Ausfälle kleiner als $c + 1$, dann soll der Posten angenommen werden, andernfalls soll er abgelehnt werden. Zu vorgegebener »akzeptierbarer mittlerer Lebensdauer« μ_{1-a} (a. m. l.) und zu vorgegebener »nicht akzeptierbarer mittlerer Lebensdauer« μ_b (u. m. l.) muß nun die Anzahl n bzw. die Schranke c geeignet gewählt werden:

Bei $E(t) = \mu \leq \mu_b$ darf die Annahmewahrscheinlichkeit (OC) $\mathbf{P}\{z \leq c\}$ nicht größer als b bzw. bei $\mu \geq \mu_{1-a}$ nicht kleiner als a sein.
Da die Annahmewahrscheinlichkeit eine Funktion

$$L(p[\lambda]) = \sum_{z=1}^{c} \binom{n}{z} (p[\lambda])^z (1 - p[\lambda])^{n-z}$$

$$\left(p[\lambda] = 1 - \sum_{i=1}^{3} p_i R(\mu_i, \beta_i)\ ;\ R(\mu, \beta) = \exp\left[-\left(\frac{T}{\mu}\Gamma\left(1 + \frac{1}{\beta}\right)\right)^{\beta}\right]\right)$$

des Vektors $\lambda = (p_1, p_2, p_3, \mu_1, \mu_2, \mu_3, \beta_1, \beta_2, \beta_3)$ der unbekannten Parameter der Verteilung ist, müssen n und c aus den Gleichungen

$$b = \max_{\lambda \in \Lambda_b} L(p[\lambda]) = L(\min_{\lambda \in \Lambda_b} p[\lambda]) = L(p_b) \tag{1}$$

$$1 - a = \min_{\lambda \in \Lambda_a} L(p[\lambda]) = L(\max_{\lambda \in \Lambda_a} p[\lambda]) = L(p_{1-a})$$

bestimmt werden. Hierbei ist Λ_b die Menge der Vektoren λ, welche die Bedingungen

$$\sum_{i=1}^{3} p_i \mu_i \leq \mu_t$$

$$p_i \geq 0\,,\ \sum_{i=1}^{3} p_i = 1$$

erfüllen, bzw. ist Λ_a die Menge der Vektoren λ, welche die Bedingungen

$$\sum_{i=1}^{3} p_i \mu_i \geq \mu_{1-a}$$

$$p_i \geq 0\,,\ \sum_{i=1}^{3} p_i = 1$$

erfüllen.
Wir setzen voraus, daß

$$\underline{\mu}_1 \leq \bar{\mu}_1 \leq \underline{\mu}_2 \leq \bar{\mu}_2 \leq \underline{\mu}_3 \leq \bar{\mu}_3$$

$$0{,}01 < \underline{\beta} \leq \bar{\beta} < 10 \tag{2}$$

$$\underline{\mu}_1 \leq \mu_b < \mu_{1-a} \leq \bar{\mu}_3$$

gilt. Für die Prüfdauer fordern wir

$$T \leq T_* = \min \left\{ \underline{\mu}_1 ; \frac{\underline{\mu}_1}{\Gamma\left(1+\frac{1}{\underline{\beta}}\right)} \left(\frac{\underline{\beta}+1}{\underline{\beta}}\right)^{\frac{1}{\underline{\beta}}} ; \frac{\bar{\mu}_1}{\Gamma\left(1+\frac{1}{\bar{\beta}}\right)} \left(\frac{\bar{\beta}+1}{\bar{\beta}}\right)^{\frac{1}{\bar{\beta}}} \right\}. \tag{3}$$

Diese Bedingung bedeutet in der Praxis meist keine Einschränkung, sie ist beispielsweise erfüllt, wenn $1 \leq \underline{\beta}$ und $T \leq \underline{\mu}_1$ ist. Die Berechnung der Extrema in (1) ergibt dann:

$$p_{1-a} = 1 - \frac{\bar{\mu}_3 - \mu_{1-a}}{\bar{\mu}_3 - \underline{\mu}_1} R(\underline{\mu}_1, \underline{\beta}) - \frac{\mu_{1-a} - \underline{\mu}_1}{\bar{\mu}_3 - \underline{\mu}_1} R(\bar{\mu}_3, \underline{\beta})$$

und

$$p_b = 1 - R(\mu_b, \bar{\beta}) \qquad (\text{bei } \mu_b \in [\underline{\mu}_i, \bar{\mu}_i],\ i = 1 \text{ oder } 2 \text{ oder } 3)$$

bzw.

$$p_b = 1 - \frac{\underline{\mu}_2 - \mu_b}{\underline{\mu}_2 - \bar{\mu}_1} R(\bar{\mu}_1, \bar{\beta}) - \frac{\mu_b - \bar{\mu}_1}{\underline{\mu}_2 - \bar{\mu}_1} R(\underline{\mu}_2, \bar{\beta}) \qquad (\text{bei } \bar{\mu}_1 \leq \mu_b \leq \underline{\mu}_2)$$

bzw.

$$p_b = 1 - \frac{\underline{\mu}_3 - \mu_b}{\underline{\mu}_3 - \bar{\mu}_2} R(\bar{\mu}_2, \bar{\beta}) - \frac{\mu_b - \bar{\mu}_2}{\underline{\mu}_3 - \bar{\mu}_2} R(\underline{\mu}_3, \bar{\beta}) \qquad (\text{bei } \bar{\mu}_2 \leq \mu_b \leq \underline{\mu}_3).$$

n und c können jetzt aus Tabellen für die Annahmeprüfung [9, 10] bestimmt werden, wobei p_b als nicht akzeptierbarer Ausschußprozentsatz (LTPD) bzw. p_{1-a} als akzeptierbarer Ausschußprozentsatz (AQL) vorgegeben sind.

3.

Wird für die Verteilung der Lebensdauer eine Mischung von zwei Weibull-Verteilungen

$$p_1 p(t|\, \mu_1 \beta_1) + p_3 p(t|\, \mu_3 \beta_3) \qquad (p_i \geq 0,\ p_1 + p_3 = 1)$$

angesetzt, dann können wir unter den Voraussetzungen (2) und (3) n und c wie in 2. bestimmen, wenn wir jetzt setzen:

$$p_{1-a} = 1 - \frac{\bar{\mu}_3 - \mu_{1-a}}{\bar{\mu}_3 - \underline{\mu}_1} R(\underline{\mu}_1, \underline{\beta}) - \frac{\mu_{1-a} - \underline{\mu}_1}{\bar{\mu}_3 - \underline{\mu}_1} R(\bar{\mu}_3, \underline{\beta})$$

und

$$p_b = 1 - R(\mu_b, \bar{\beta}) \qquad (\text{bei } \mu_b \in [\underline{\mu}_i, \bar{\mu}_i],\ i = 1 \text{ oder } 3)$$

bzw. bei $\bar{\mu}_1 \leq \mu_b \leq \underline{\mu}_3$:

$$p_b = 1 - \frac{\underline{\mu}_3 - \mu_b}{\underline{\mu}_3 - \bar{\mu}_1} R(\bar{\mu}_1, \bar{\beta}) - \frac{\mu_b - \bar{\mu}_1}{\underline{\mu}_3 - \bar{\mu}_1} R(\underline{\mu}_3, \bar{\beta})$$

4.

Wird für die Verteilung der Lebensdauer eine Weibull-Verteilung $p(t|\mu, \beta)$ mit unbekannten Parametern μ und β angesetzt und ist bekannt, daß β in $[\underline{\beta}, \bar{\beta}]$ liegt, dann können wir unter der Voraussetzung

$$T \leq \underline{\mu}_1 \leq \mu \,;\; \mu_1 \leq \mu_b < \mu_{1-a}$$

n und c wieder wie in 2. berechnen, wenn wir jetzt setzen:

$$p_{1-a} = 1 - R(\mu_{1-a}, \underline{\beta})$$

und (4)

$$p_b = 1 - R(\mu_b, \bar{\beta}) \,.$$

Würde man den Gestaltsparameter durch einen Wert $\hat{\beta} \in [\underline{\beta}, \bar{\beta}]$ schätzen und die Tabellen von Goode und Kao verwenden, dann ergäbe sich, wie man leicht nachprüft, eine kleinere Anzahl n von Elementen. Um aber das vorgegebene Produzentenrisiko a bzw. das Konsumentenrisiko b auch bei von $\tilde{\beta}$ abweichenden Parametern nicht zu überschreiten, muß die durch (4) gegebene Erhöhung der Anzahl n in Kauf genommen werden.

5.

Abschließend machen wir einige Bemerkungen zu den angegebenen Resultaten.

(a) Der Fall einer Mischung von Exponentialverteilungen mit verschiedenen mittleren Lebensdauern ergibt sich bei $\underline{\beta} = \bar{\beta} = 1$.

(b) Die Formeln aus 2. und 3. können auch angewendet werden, wenn die bei der Mischung eingehenden Verteilungen einen gemeinsamen und bekannten Gestaltsparameter β oder bekannte Erwartungswerte besitzen. Dann muß $\beta = \underline{\beta} = \bar{\beta}$ bzw.

$$\underline{\mu}_i = \mu_i = \bar{\mu}_i$$

gesetzt werden.

(c) Ist für jeden Gestaltsparameter β_i ein Intervall

$$[\underline{\beta}_i, \bar{\beta}_i]$$

vorgegeben, dann kann $\beta = \min_i \underline{\beta}_i$ und $\bar{\beta} = \max_i \bar{\beta}_i$ gesetzt werden und n bzw. c wie oben bestimmt werden. Die vorgegebenen Risikos werden natürlich nicht überschritten, andererseits könnte bei einer Berücksichtigung der verschiedenen Intervallgrenzen auf anderem Wege eventuell ein kleineres n bestimmt werden.

(d) Indem für verschiedene Werte T die entsprechenden Anzahlen $n(T)$ berechnet werden, kann auf Grund der Kurve $n(T)$ ein ökonomisch günstiger Versuchsplan gewählt werden. Muß die Anzahl n aus technischen oder öko-

nomischen Gründen klein gehalten werden, dann kann eine übermäßig lange Prüfzeit notwendig werden. In diesem Fall wäre eine Kontrolle auf Grund der beobachteten Lebensdauern der geprüften Elementen angebracht. In weiteren Arbeiten wollen wir uns mit der Konstruktion solcher, mit gewissen analytischen Schwierigkeiten verbundenen, Verfahren beschäftigen.

LITERATUR

1. Шор, Я. Б. Статистикеские методы анализа и контроля качества и надежности, Москва (1962).
2. Weibull, W., Fatigue Testing and the Analysis of Results. (1961).
3. Semi-conductor Reliability. Elizabeth (1962).
4. Goode, H. P., Kao, J. H. K., Sampling Plans Based on the Weibull Distribution. *Proc. 7. Nat. Symp. on Reliability and Quality Control* (1961).
5. Goode, H. P., Kao, J. H., K. An Adaptation of the MTL-STD-105B Plan to Reliability and Life Testing Application. *Trans. 15. Ann. Convention Am. Soc. Quality Control* (1961).
6. Goode, H. P., Kao, J. H. K., Quality Control and Reliability Technical Report TR 3, Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (Mean Life Criterion). U. S. Gov. Printing Office (1961).
7. Gottfried, P., Roberts, H., Some Pitfalls of the Weibull-Distribution. *Proc. 9. Nat. Symp. on Reliability and Quality Control* (1963).
8. Kao, J. H. K., *Trans. Am. Soc. Quality Control* (1959).
9. Grubbs, T. E., On Designing Simple Sampling Inspection Plans. *Ann. Math. Stat.* **20,** (1949).
10. Military Standard 105, Sampling Procedures and Tables for Inspection by Attributes. Gov. Print. Office.
11. Are Statistical Life Testing Procedures Robust? *Nat. Bureau of Standards, Industrial Quality Control* **17,** (1961).

ON THE NUMBER OF INTERSECTIONS BETWEEN TWO EMPIRICAL DISTRIBUTION FUNCTIONS

by

E. CSÁKI

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary

1. INTRODUCTION AND NOTATIONS

Let $(\xi_1, \ldots, \xi_n)$ and $(\eta_1, \ldots, \eta_n)$ be two samples of equal sizes. Let us denote their distribution functions by $F(x)$ and $G(x)$ and their empirical distribution functions by $F_n(x)$ and $G_n(x)$. Among other properties the number of intersections of the two empirical distribution functions has been considered in [1] and [2], provided that the hypothesis $F(x) = G(x)$ is true. In the present paper we shall consider problems in connection with the number of intersections in case of some alternative.

We shall make use of the well-known random walk model concerning the two-sample problem. Let us arrange the two samples in one combined sample according to the order of magnitude: $\zeta_1 < \zeta_2 < \ldots < \zeta_{2n}$. Throughout this paper $F(x)$ and $G(x)$ will be continuous, consequently no equal elements will occur among the sample elements with probability 1. Let us introduce the random variables ϑ_i

$$\vartheta_i = \begin{cases} +1 & \text{if} \quad \zeta_i = \eta_j \quad \text{for some } j \\ -1 & \text{if} \quad \zeta_i = \xi_k \quad \text{for some } k \end{cases}$$

and put $s_0 = 0$, $s_i = \vartheta_1 + \vartheta_2 + \ldots + \vartheta_i$ $(i = 1, 2, \ldots, 2n)$.

The number of intersections λ between the two empirical distribution functions $F_n(x)$ and $G_n(x)$ equals the number of indices i such that $s_{i-1}\, s_{i+1} = -1$.

We shall consider the symmetric random walk, i.e. the independent, identically distributed random variables ε_i $(i = 1, 2, \ldots)$ with $\mathbf{P}(\varepsilon_i = +1) = \mathbf{P}(\varepsilon_i = -1) = \frac{1}{2}$ and their partial sums $s_0 = 0$, $s_i = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_i (i = 1, 2, \ldots)$. Let λ_r be the change of signs in the sequence $(s_0, s_1, \ldots, s_r)$, i.e. the number of indices i for which $s_{i-1}\, s_{i+1} = -1$ and $i < r$. Our consideration is based on the formulae (see [2] and [3]):

$$\mathbf{P}(\lambda_{2r} \geq l \mid s_{2r} = 0) = 2\,\frac{\binom{2r-1}{r+l}}{\binom{2r}{r}} \qquad l = 0, 1, \ldots, r-1 \tag{1}$$

$$\mathbf{P}(\lambda_{2r} \geq l \,|\, s_{2r} = 2k) = \frac{\binom{2r}{r+k+l}}{\binom{2r}{r+k}} \qquad \begin{matrix} l = 0, 1, \ldots r-k \\ k > 0 \end{matrix} \tag{2}$$

$$\mathbf{P}(\lambda_{2r-1} \geq l \,|\, s_{2r-1} = 2k-1) = \frac{\binom{2r-1}{r+k+l-1}}{\binom{2r-1}{r+k-1}} \qquad \begin{matrix} l = 0, 1, \ldots, r-k \\ k > 0 \end{matrix} \tag{3}$$

2. A SIMPLE ALTERNATIVE

The exact distribution of λ for the alternative situation

$$\begin{aligned} F(x) &= x \qquad 0 \leq x \leq 1 \\ G(x) &= \frac{x}{z} \qquad 0 \leq x \leq z, \end{aligned} \tag{4}$$

where z is a constant $(0 < z \leq 1)$, can be given by

THEOREM 1. *If $F(x)$ and $G(x)$ is given by* (4), *then*

$$\mathbf{P}(\lambda \geq l) = \frac{1}{\binom{2n}{n}} \left[\binom{2n}{n+l} z^l - \frac{l}{n} \binom{2n}{n+l} z^n \right], \quad l = 0, 1, \ldots, n-1. \tag{5}$$

PROOF. According to our assumption the whole sample $(\eta_1, \eta_2, \ldots, \eta_n)$ lies within the interval $(0, z)$. Let us denote by ν the number of ξ-s lying in the interval $(0, z)$. Then using the formulae (1), (2), (3) we obtain

$$\begin{aligned} \mathbf{P}(\lambda \geq l) &= \sum_{j=0}^{n} \mathbf{P}(\lambda \geq l \,|\, \nu = j)\, \mathbf{P}(\nu = j) = \\ &= \sum_{j=0}^{n} \mathbf{P}(\lambda_{n+j} \geq l \,|\, s_{n+j} = n - j) \binom{n}{j} z^j (1-z)^{n-j} = \\ &= 2 \frac{\binom{2n-1}{n+l}}{\binom{2n}{n}} z^n + \sum_{j=l}^{n-1} \frac{\binom{n+j}{n+l}}{\binom{n+j}{n}} \binom{n}{j} z^j (1-z)^{n-j} = \\ &= 2 \frac{\binom{2n-1}{n+l}}{\binom{2n}{n}} z^n + \frac{\binom{2n}{n+l}}{\binom{2n}{n}} \sum_{j=l}^{n-1} \binom{n-l}{j-l} z^j (1-z)^{n-j} = \end{aligned}$$

$$= 2\frac{\binom{2n-1}{n+l}}{\binom{2n}{n}} z^n + \frac{\binom{2n}{n+l}}{\binom{2n}{n}} z^l(1-z^{n-l}),$$

which is equivalent to (5).

It should be noted that the distribution of λ is obviously invariant under monotone and continuous mapping of the random variables, applying the same transformation to each ξ_i and η_j.

3. ASYMPTOTIC FORMULAE FOR THE CONDITIONAL DISTRIBUTION OF λ_N IN THE CASE OF SYMMETRIC RANDOM WALK

For general alternatives the determination of the distribution of λ seems to be complicated. Therefore we are satisfied with the determination of asymptotic distribution of λ for a piecewise linear alternative. Before doing so we shall give some asymptotic formulae for the conditional distribution of λ_N in the symmetric random walk. These results were obtained by I. Vincze and the author [4]. The distribution of λ_N under the condition $s_N = f(N)$ will be considered, $f(N)$ being an integer-valued function of N. Four different cases depending on the order of $f(N)$ are to be distinguished. Making use of the formulae (2) and (3) and the Stirling formula we can get the following limiting distributions:

THEOREM 2.

(a) *if* $f(N) = o(\sqrt{N})$,

then

$$\lim_{N\to\infty} \mathbf{P}\left(\lambda_N < y\sqrt{N} \mid s_N = f(N)\right) = 1 - \mathrm{e}^{-2y^2} \quad y \geq 0 \tag{6}$$

(b) *if* $f(N) \sim c\sqrt{N}$,

then

$$\lim_{N\to\infty} \mathbf{P}\left(\lambda_N < y\sqrt{N} \mid s_N = f(N)\right) = 1 - \mathrm{e}^{-2(y+c)^2+2c^2} \quad y \geq 0 \tag{7}$$

(c) *if* $f(N) = o(N)$, $\sqrt{N} = o(f(N))$

and $g(N)$ *is a function such that*

$$\lim_{N\to\infty} \frac{f(N)\, g(N)}{N} = C > 0,$$

then

$$\lim_{N\to\infty} \mathbf{P}\left(\lambda_N < y\, g(N)\right) = 1 - \mathrm{e}^{-2cy} \quad y \geq 0 \tag{8}$$

(d) *if* $f(N) \sim cN$,

then

$$\lim_{N\to\infty} \mathbf{P}\big(\lambda_N = l \,|\, s_N = f(N)\big) = \frac{2c}{1+c}\left(\frac{1-c}{1+c}\right)^l \qquad l = 0, 1, 3, \ldots \tag{9}$$

i.e. in the latter case the number of intersections is finite with probability one.

4. SOME OTHER ALTERNATIVES

The following alternative will now be considered:

$$F(x) = x \qquad 0 \le x \le 1$$

$$G(x) = \begin{cases} gx & 0 \le x \le z \\ \dfrac{1-gz}{1-z}\,x + (g-1)\dfrac{z}{1-z} & z \le x \le 1, \end{cases} \tag{10}$$

where $0 < z < 1$ and $1 < g \le \dfrac{1}{z}$.

Denoting by ν and μ the number of elements lying in the interval $(0, z)$ from the samples $(\xi_1, \xi_2, \ldots, \xi_n)$ and $(\eta_1, \eta_2, \ldots, \eta_n)$ resp., we obtain

$$\mathbf{P}(\lambda \ge l) = \sum_{(j,k)} \mathbf{P}(\lambda \ge l \,|\, \nu = j,\ \mu = k)\,\mathbf{P}(\nu = j,\ \mu = k) = \tag{11}$$

$$= \sum_{(j,k)} \mathbf{P}(\lambda_{2n} \ge l \,|\, s_{j+k} = k - j,\ s_{2n} = 0) \binom{n}{k} (gz)^k (1-gz)^{n-k} \binom{n}{j} z^j (1-z)^{n-j}.$$

The probabilities $\mathbf{P}(\lambda_{2n} \ge l \,|\, s_{j+k} = k - j,\ s_{2n} = 0)$ can be determined explicitly as λ_{2n} may be considered as a sum

$$\lambda_{2n} = \lambda_{j+k} + \lambda_{2n-j-k}$$

where λ_{j+k} is the number of changes of signs in the sequence $(s_0, s_1, \ldots, s_{j+k})$, while λ_{2n-j-k} is that of $(s_{j+k+1}, \ldots, s_{2n})$ and they are independent under the condition $s_{j+k} = k - j$, $s_{2n} = 0$. The formula however is rather complicated. Instead of this we consider the asymptotic distribution of λ, when $G(x)$ depends on n, in such a way that z is fixed and

$$g = 1 + \frac{2c}{(2n)^\alpha} \tag{12}$$

where $c > 0$ and $0 \le \alpha < \dfrac{1}{2}$.

THEOREM 3. *If $F(x)$ and $G(x)$ are given by* (10) *and* (12) *with z fixed, then*

(a) for $0 < \alpha < \dfrac{1}{2}$

$$\lim_{n\to\infty} \mathbf{P}\big(\lambda \ge y(2n)^\alpha\big) = \begin{cases} \dfrac{z}{2z-1}\,\mathrm{e}^{-2cy} - \dfrac{1-z}{2z-1}\,\mathrm{e}^{-\frac{2cz}{1-z}y}, & \text{if } z \ne \dfrac{1}{2},\ y \ge 0 \\ \mathrm{e}^{-2cy}(1+2cy), & \text{if } z = \dfrac{1}{2},\ y \ge 0 \end{cases} \tag{13}$$

(b) *for* $\alpha = 0$

$$\lim_{n\to\infty} \mathbf{P}(\lambda = l) =$$

$$= \begin{cases} \dfrac{2cz}{2cz+2z-1}\left[\left(\dfrac{1}{1+2c}\right)^{l+1} - \left(\dfrac{1-z-2cz}{1-z}\right)^{l+1}\right] & \text{if } z(1+c) \neq \dfrac{1}{2},\ l = 0,1,2,\ldots \\ (l+1)\left(\dfrac{2c}{1+2c}\right)^2\left(\dfrac{1}{1+2c}\right)^l & \text{if } z(1+c) = \dfrac{1}{2},\ l = 0,1,2,\ldots \end{cases} \tag{14}$$

PROOF. The summation in (11) can be divided into two parts, the first part being summed up over j and k such that $j \sim nz$ and $k \sim ngz$, the second part over the remaining j and k. The latter can be neglected by virtue of the normal approximation for binomial probabilities. For the first part $j \sim nz$, $k \sim ngz$, thus if $0 < \alpha < \dfrac{1}{2}$,

$$j + k \sim nz(1+g) \sim 2nz$$

$$k - j \sim nz(g-1) = cz(2n)^{1-\alpha}.$$

Put $N_1 = [2nz]$ and $N_2 = [2n(1-z)]$, then

$$k - j \sim cz(2n)^{1-\alpha} \sim cz^{\alpha}N_1^{1-\alpha}$$

and

$$k - j \sim c\frac{z}{1-z}(1-z)^{\alpha}N_2^{1-\alpha}.$$

Using the limiting distributions (8) in Theorem 2, we get

$$\mathbf{P}(\lambda_{N_1} \geq y(2n)^{\alpha} \mid s_{N_1} \sim cz^{\alpha}N_1^{1-\alpha}) \sim e^{-2cy}$$

and

$$\mathbf{P}(\lambda_{N_2} \geq y(2n)^{\alpha} \mid s_{N_2} \sim c\frac{z}{1-z}(1-z)^{\alpha}N_2^{1-\alpha}) \sim e^{-\frac{2cz}{1-z}y}.$$

In case $\alpha = 0 : k + j \sim 2nz(1+c)$, and $k - j \sim 2nzc$, put $N_1 = [2nz(1+c)]$, $N_2 = [2n(1-z-cz)]$, then by (9)

$$\mathbf{P}\left(\lambda_{N_1} = l \mid s_{N_1} \sim \frac{c}{1+c}N_1\right) \sim \frac{2c}{1+2c}\left(\frac{1}{1+2c}\right)^l$$

and

$$\mathbf{P}(\lambda_{N_2} = l \mid s_{N_2} \sim \frac{cz}{1-z-cz}N_2) \sim \frac{2cz}{1-z}\left(\frac{1-z-2cz}{1-z}\right)^l.$$

Thus the conditional distribution of $\frac{\lambda}{(2n)^\alpha} = \frac{\lambda_{N_1}}{(2n)^\alpha} + \frac{\lambda_{N_2}}{(2n)^\alpha}$ is asymptotically equal to the convolution $H(y)$ of the distributions given above. Therefore

$$\mathbf{P}(\lambda \geq y(2n)^\alpha) \sim \sum_{\substack{j \sim nz \\ k \sim ngz}} H(y) \binom{n}{k} (gz)^k (1 - gz)^{n-k} \binom{n}{j} z^j (1 - z)^{n-j} \sim H(y).$$

which is given in (13) and (14).

We should remark that similar considerations can be carried out for other piecewise linear alternatives.

REFERENCES

1. Михалевич, В. С. О взаимном расположении двух эмпирических функций распределения, (On the mutual position of two empirical distributions.) ДАН СССР **85,** *3* 485—488 (1952)
2. Csáki, E., Vincze, I., On Some Problems Connected with the Galton Test. *Publ. Math. Inst. Hung. Acad. Sci.* **6,** 97–109 (1961)
3. Kanwar Sen, On Some Combinatorial Relations Concerning the Symmetric Random Walk. *Publ. Math. Inst. Hung. Acad. Sci.* **9,** 335–357 (1964).
4. Vincze, I., On Some Questions Connected with Two Sample Tests of Smirnov-type. *Proc. Fifth Berkeley Symp. Prob. Stat.* (to be published).

DIE ANWENDUNG DER C-KARTE ZUR KONTROLLE VON ZIEHTEILEN AUS FEINBLECHEN

von

B. Czyżewski

H. Cegielski Werke
Poznan, Polen

Ziehbearbeitung ist charakterisiert durch eine große Mäßigkeit der Produkte in Beziehung auf Form und Masse. Gleichzeitig ist fehlerhafte Produktion meistens sehr nieder. Das schafft einige Schwierigkeiten bei der SQK-Methode weil:

(a) die große Anzahl von Massen und Formen gezogener Produkte die Anwendung der alternativen SQK Methoden beschränkt;

(b) der Stichprobenumfang bei der alternativen SQK-Methode mit dem Ausschußanteil nach der Formel: $n\bar{p} = \bar{c}/n$ — der Stichprobenumfang, $\bar{p}$ — durchschnittlicher Ausschußanteil in der Produktion, c — durchschnittliche Anzahl der fehlerhaftigen Stücke in den Stichproben verbunden ist. Daraus ergibt sich, daß bei konstantem c-Wert der Stichprobenumfang umso größer wird je kleiner der Ausschußanteil ist;

(c) mit größerem Stichprobenumfang vergrößert sich der Zeitaufwand für die Kontrolle und bei zerstörenden Prüfmethoden vergrößert sich der Verlust an Probestücken.

Vor der Einführung der SQK gab als Hauptaufgabe die Beschränkung des Stichprobenumfangs auf eine ökonomische Größe, im Hinblick auf die Prüfzeit, eventuell auf die Anzahl zu zerstörender Proben, bei genügender Empfindlichkeit der Kontrollkarte.

Im Falle, den ich schildern möchte, wurde die »c-Karte« als Kontrollkarte für die Zahl der fehlerhaften Stücke eingeführt. Die Empfindlichkeit dieser Karte in Bezug auf die Schwankungen im technologischen Prozess wird durch die Erfüllung der Bedingung: $c \geq 1$ gesichert. Die Größe des Stichprobenumfangs und der Ausschußanteil der Produktion geben eine hyperbolische Abhängigkeit $n\bar{p} = \bar{c}$, welche für $\bar{c} = 1$ in Abb. 1 gezeigt wird. Aus Abb. 1 ergibt sich, daß der Stichprobenumfang bei Ausschußanteil $\bar{p} < 0{,}05$ stark zu wachsen anfängt. Für diese Ausschußanteile wurde der Stichprobenumfang mit $n \leq 25$ begrenzt, wodurch man $c < 1$ erhielt. Es ergab sich, daß bei einer sehr scharfen Klassifikation aller Merkmale der Produkten, die Empfindlichkeit der Karte sich unbedeutend senkte.

Bei dem Ausschußanteil $\bar{p} \leq 0{,}05$ unterliegt die Zahl der fehlerhaften Stücke einer Poissonischen Verteilung, so daß man die Kontrollgrenze aus der nachstehenden Formel berechnen kann:

$$\mathrm{OKL}_c = \bar{c} + t\sqrt{\bar{c}} \tag{1}$$

(meistens $t = 3$).

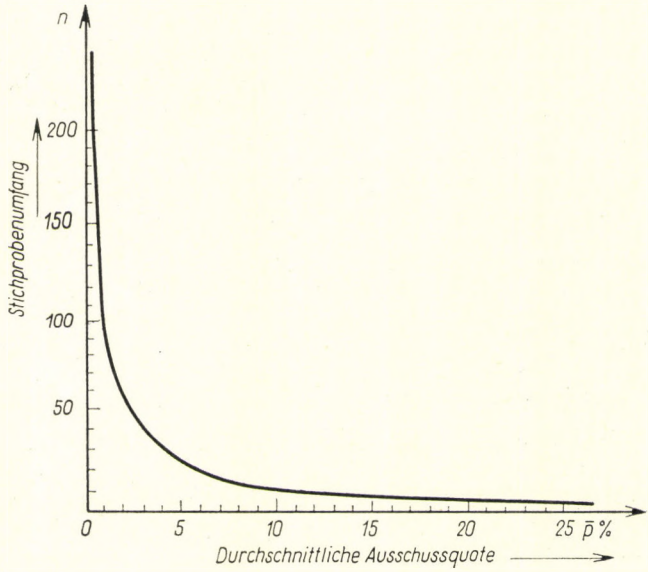


ABB. 1

Abhängigkeit des Stichprobenumfangs von der durchschnittlichen Ausschußquote $\bar{p}$ unter Voraussetzung daß $n\bar{p} = 1$ für alternative SQK-Methoden

Die Begrenzung der Zahl der fehlerhaften Stücke nach unten bei der Probe im Falle $\bar{c} < t^2$ ist zwecklos, weil in diesem Falle $\mathrm{UKL}_c = 0$ gesetzt werden kann ($\mathrm{UKL}_c < 0$ wäre sinnlos).

Die Wahrscheinlichkeit, daß sich in der Stichprobe eine Anzahl fehlerhafter Stücke befindet, welche die obere Kontrollgrenze überschreitet (errechnet aus Formel 1), ist von der Formel $n\bar{p} = \bar{c}$ abhängig und deshalb ist der Wert $t = 3$ nicht genau gleich, 0,27% wie es bei der Normalverteilung der Fall ist.

Deshalb auch, wenn wir Wert darauf legen, daß die Kontrollkarten unabhängig von dem Wert des Produktes np mit der gleichen Wahrscheinlichkeit der Überschreitung der Kontrollgrenzen funktionieren, muß für verschiedene c-Werte die Wahrscheinlichkeit dafür, daß in der Stichprobe $c = 0, 1, 2, \ldots$ fehlerhafte Stücke sich befinden, aus der Poissonschen Formel ausgewertet werden. Aus den erhaltenen Werten können wir dann durch Summieren Tabelle 1 zusammenstellen, die die Wahrscheinlichkeit dafür, daß in der Stichprobe höchstens 0, 1, 2, 3, ... fehlerhafte Stücke sich befinden, angibt.

Der Unterschied zwischen 1 (Wahrscheinlichkeit des Erhaltens von $c = 0, 1, 2, \ldots$ fehlerhaften Stücke in der Stichprobe) und den erhaltenen Werten ist die Wahrscheinlichkeit, daß mehr als 0, 1, 2, 3, ... fehlerhafte Stücke in der Stichprobe erhalten werden.

Auf Grund von Tabelle I wurde das Diagram 2 angefertigt.

Tabelle I

Die Wahrscheinlichkeit des Vorkommens von nicht mehr als c fehlerhaften Stücken bei der Stichprobe

c \ z	0	1	2	3	4	5	6	7	8	9	10
0,05	0,9512	0,9988	1,0000								
0,10	0,9048	0,9953	0,9998	1,0000							
0,15	0,8607	0,9898	0,9995	1,0000			$P_c = \frac{\bar{c}^c}{c!} \cdot e^{-\bar{c}}$				
0,20	0,8187	0,9624	0,9988	0,9999	1,0000						
0,25	0,7788	0,9735	0,9978	0,9998	1,0000		$P = P_0 + P_1 + \ldots + P_c$				
0,30	0,7408	0,9630	0,9964	0,9997	1,0000						
0,40	0,6703	0,9384	0,9920	0,9992	0,9999	1,0000					
0,50	0,6065	0,9098	0,9856	0,9982	0,9998	1,0000					
0,60	0,5488	0,8781	0,9769	0,9966	0,9996	1,0000					
0,70	0,4966	0,8442	0,9659	0,9942	0,9992	0,9999	1,0000				
0,75	0,4724	0,8267	0,9596	0,9918	0,9989	0,9998	1,0000				
0,80	0,4493	0,8088	0,9526	0,9909	0,9986	0,9998	1,0000				
0,90	0,4066	0,7725	0,9371	0,9865	0,9977	0,9997	1,0000				
1,00	0,3679	0,7358	0,9197	0,9810	0,9963	0,9994	0,9999	1,0000			
1,10	0,3329	0,6990	0,9004	0,9742	0,9945	0,9990	0,9998	0,9999	1,0000		
1,20	0,3012	0,6626	0,8795	0,9662	0,9922	0,9986	0,9997	0,9999	1,0000		
1,30	0,2725	0,6268	0,8571	0,9569	0,9893	0,9977	0,9995	0,9998	1,0000		
1,40	0,2466	0,5918	0,8335	0,9463	0,9857	0,9968	0,9994	0,9998	1,0000		
1,50	0,2231	0,5578	0,8088	0,9343	0,9814	0,9955	0,9990	0,9998	0,9999	1,0000	
1,60	0,2019	0,5249	0,7833	0,9211	0,9763	0,9939	0,9986	0,9997	0,9999	1,0000	
1,70	0,1827	0,4932	0,7572	0,9068	0,9704	0,9920	0,9981	0,9996	0,9999	1,0000	
1,80	0,1653	0,4620	0,7306	0,8913	0,9636	0,9896	0,9974	0,9994	0,9999	1,0000	
1,90	0,1496	0,4338	0,7037	0,8747	0,9559	0,9868	0,9966	0,9992	0,9998	0,9999	1,0000
2,00	0,1353	0,4060	0,6767	0,8571	0,9473	0,9834	0,9954	0,9989	0,9998	0,9999	1,0000

Tabelle II

Werte der oberen Kontrollgrenze $\mathrm{OKL}_c = \bar{c} + 3\sqrt{\bar{c}}$ für die c-Karte

c	0	0,1	0,2	0,25	0,3	0,4	0,5	0,6	0,7	0,75	0,8	0,9
0	—	1,05	1,63	1,75	1,94	2,30	2,62	2,93	3,21	3,35	3,48	3,75
1	4,00	4,24	4,49	4,60	4,72	4,94	5,17	5,39	5,60	5,71	5,82	6,02
2	6,23	6,45	6,65	6,75	6,86	7,05	7,24	7,43	7,62	7,72	7,82	8,01
3	8,19	8,37	8,56	8,65	8,75	8,92	9,11	9,28	9,46	9,55	9,63	9,81
4	10,00	10,16	10,35	10,42	10,50	10,68	10,86	11,02	11,19	11,28	11,37	11,54
5	11,70	11,87	12,05	12,12	12,20	12,36	12,53	12,70	12,87	12,95	13,02	13,18

Um das Bestimmen der Kontrollgrenzen der c-Karte zu erleichtern, wurden die Diagramme auf Abb. 3 für die am meisten gebrauchten Wahrscheinlichkeiten zusammengestellt.

Tabelle I und Abb. 2 und 3 berücksichtigen vor allem kleine Werte der Produkte $n\bar{p}$, in Übereinstimmung mit dem, was bei der Wahl der Kontrollkarte vorausgesetzt wurde.

In der Praxis der H. Cegielski-Werke werden die Kontrollinien an Hand der Formel 1 festgelegt und deshalb wurde auch zwecks Vermeidung von Berechnungen Tabelle II zusammengestellt.

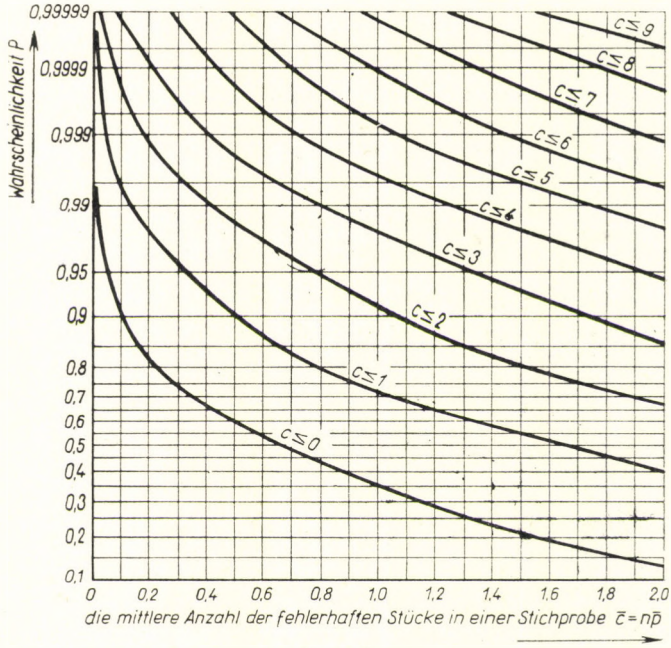


ABB. 2

Diagramm zur Berechnung der Veränderungsgrenzen der Anzahl der fehlerhaften Stücke bei der Stichprobe bei festgelegten Wahrscheinlichkeit oder zum Bestimmen der Wahrscheinlichkeit des Erhaltens von mehr als 0, 1, 2, 3, . . ., fehlerhaften Stücken in der Stichprobe

Tabelle I und die Diagramme auf Abb. 2 dienen zur Bestimmung der Wahrscheinlichkeit des Auftretens einer bestimmten Anzahl von fehlerhaften Stücken in der Stichprobe, falls ein solches Bedürfnis sich ergibt.

Die Kontrollkarte für die Zahl der fehlerhaften Stücke bei der Stichprobe unter der Voraussetzung $\bar{c} < 1$, wurde mit Erfolg in den H. Cegielski-Werken für Fahrradbeleuchtungsbestandteile in Poznan angewendet. Als Beispiel möchte ich die Vorbereitungen und die Führung der Kontrollkarte für die Fahrradlampenkörper schildern.

Auf Grund der durchgeführten Einleitungsversuche wurde der Ausschußanteil einer technologischen Ziehoperation mit $\bar{p} = 0{,}03$ festgesetzt. Es wurde weiter angenommen, daß für ein Zeitintervall von $T = 60$ Min. zwischen zwei aufeinander folgenden Stichprobenentnahmen, der aus ökonomischen Gründen beste Stichprobenempfang $n = 25$ ist; daraus folgt $c = n\bar{p} = 25.0{,}03 = 0{,}75$. Für diesen Wert ist nach Formel 1 $\mathrm{OKL}_c = 3{,}35$. An Hand der Tabelle 1 oder der Abb. 2 stellen wir fest, daß die Wahrscheinlichkeit der Überschreitung der Kontrollgrenze $\sim$ 0,004 beträgt. Das heißt daß wir nur in einem von 250 Fällen in der Stichprobe mehr als 3,35 fehlerhafte Stücke vorfinden. Wir nehmen an, daß dies genügt, um ein entsprechendes Güteniveau zu sichern. Die Werte $\bar{c} = 0{,}75$ und $\mathrm{OKL}_c =$

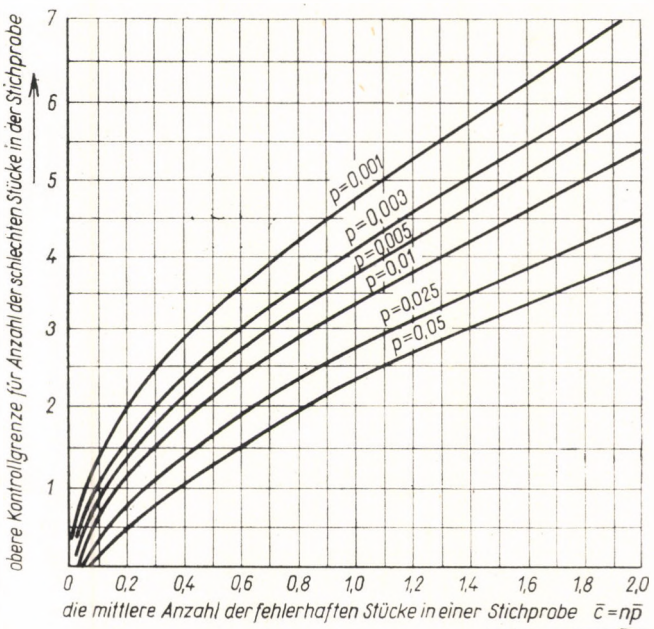


ABB. 3

Diagramm zur Berechnung der Veränderungsgrenzen der Anzahl der fehlerhaften Stücke in der Stichprobe bei verschiedenen Werten der Wahrscheinlichkeit

3,35 werden in das Diagramm der Kontrollkarte eingetragen (Abb. 4) und wir beginnen mit der laufenden Kontrolle an Hand der speziel für diese Operation bearbeitete Kontrollinstruktion (Abb. 5).

Diese Instruktion gibt die Merkmale der Werkstücke die der Kontrolle unterliegen und Methoden SQK sowie die nötigen Meßinstrumente an.

Die auf Abb. 4 dargestellte Kontrollkarte ist eine Kopie der ersten Karte, welche für diese Operation geführt wurde. Wir sehen, daß in der ersten Zeit die Kontrollinie oft überschritten wurde und wenn sie diese auch nicht übertrat, war doch die Anzahl der fehlerhaften Stücke in den Stichproben nie kleiner als 3. Eine Besserung sehen wir in der 20. Stichprobe; also nach 4 Tagen, von der 40. Stichprobe an ist mit Ausnahme der Stichproben 45, 47, 48 und 49 der Verlauf der Produktion befriedigend. Also in der Zeit von 7 Tagen gelang es den Produktionsprozeß zu stabilisieren und die Anzahl der fehlerhaften Stücke beträchtlich zu vermindern. Die Abnahme der Anzahl der fehlerhaften Stücke erfolgte bei fast allen von der SQK erfassten technologischen Operationen. Die allgemeine Anzahl fehlerhaften Stücke von Bestandteilen und fertigen Produkten senkte sich nach Einführung des SQK beträchtlich, was aus Abb. 6 und 7 ersichtlich ist.

Wie wir aus dem obigen Beispiel ersehen, kann eine systematische Kontrolle des technologischen Prozesses durch schnelles Entdecken und schnelles Beseitigen von Störungen im Produktionsprozeß zur Erhöhung der Qualität

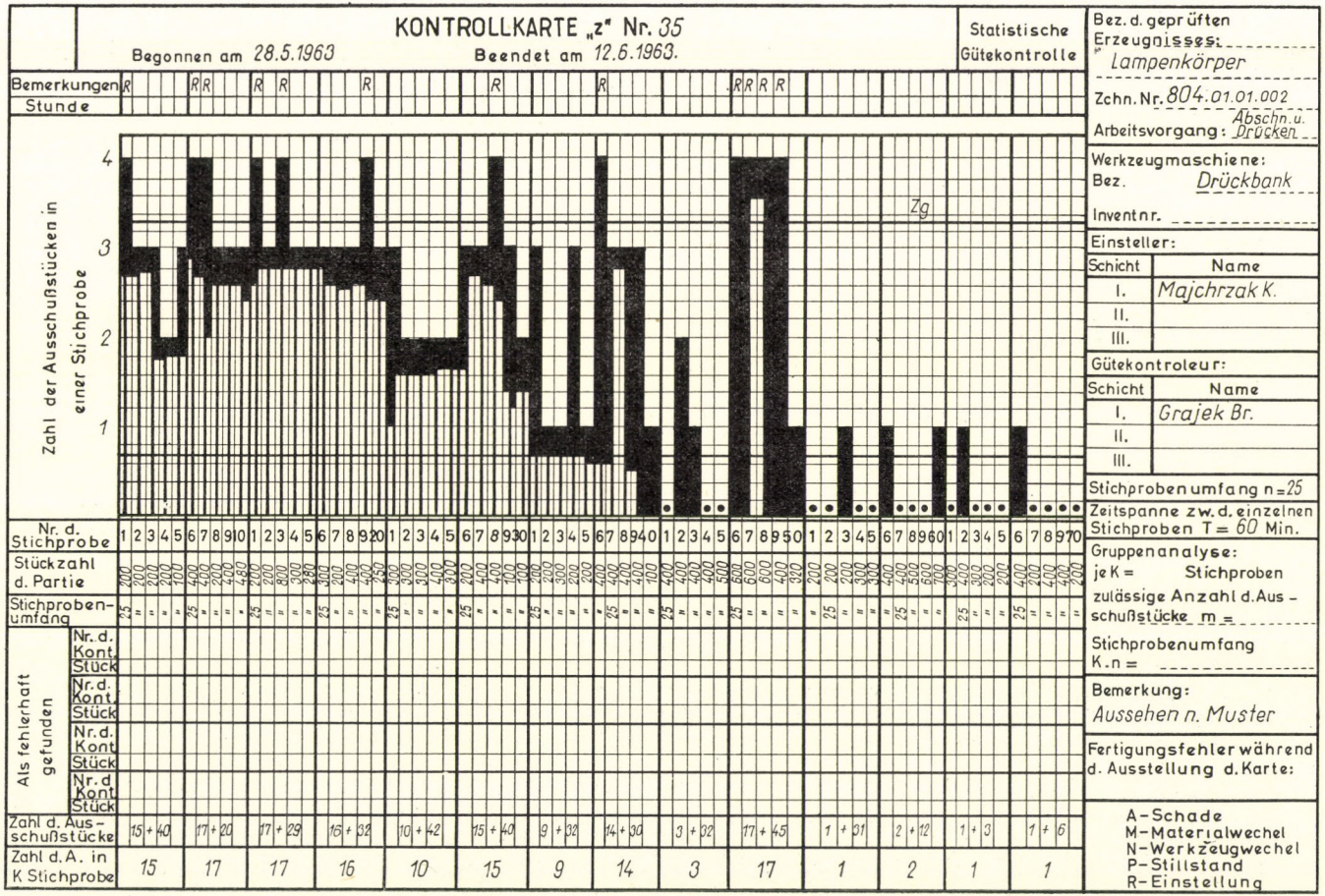


Abb. 4

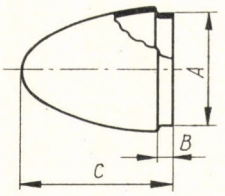
Die in den H. Cegielski-Werken angewandte Kontrollkarte für $\bar{c} < 1$ ($\bar{c} = 0{,}75$)

Kontrollanweisung — Stat. Gütekontrolle

Geprüfte Kennzeichen		Meßwerkzeug		Art der Stat. Gütekontrolle
Bezeichnung	Abmessung	Symbol	Ablesungsgenauigkeit	
A	$\phi 69{,}1 - 0{,}2$	804.01.01	002/55	
B	$4{,}5^{+0{,}5}$	804.01.01	002/57	Karte „Z"
C	$77 \pm 0{,}5$	MAU a 140		
D	Äußere Gestaltung			
E				
F				
G				
H				
I				
K				
L				
M				
N				
O				
P				
R				
S				
T				
U				

Bez. d. Bestandteils: Lampenkörper
Zchn. Nr. 804.01.01.002
Nr. u. Bez. d. Arbeitsvorganges Abschneiden u. Drücken
Umfang d. Stichprobe: n = 25 Stück
Häufigkeit d. Stichprobenahme t = 60 Min.

Skizze d. Arbeitsvorganges:



Bemerkungen: Äußere Gestaltung ist mit dem Muster zu vergleichen

ABB. 5
Kontrollinstruktion für Ziehbearbeitung

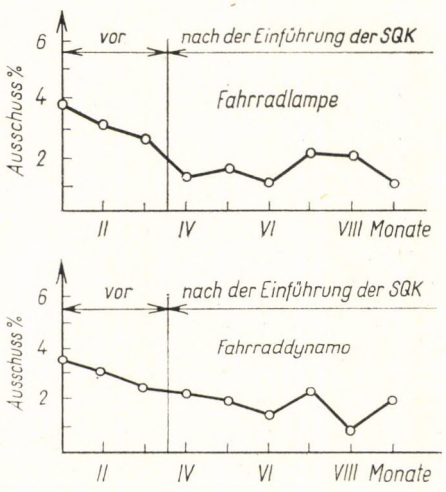


ABB. 6
Ausschußquote der Produktion vor und nach der Einführung der SQK

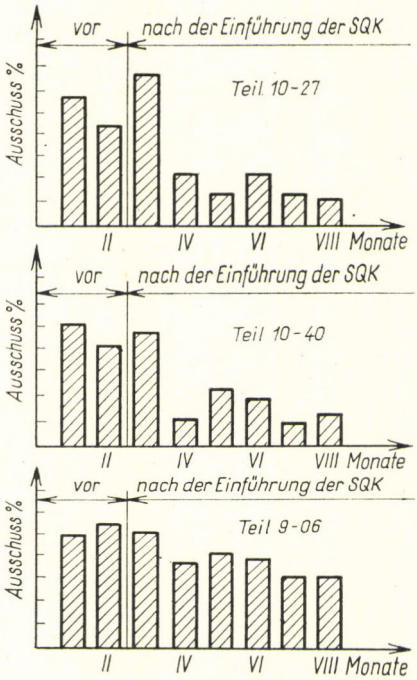


ABB. 7
Ausschußquote einiger Bestandteile vor und nach der Einführung der SQK

führen. Es ist nicht ohne Bedeutung, daß der Arbeiter oder seine Vorgesetzten zu jeder Zeit die Güte der Produktion mit der vergangenen Periode oder mit den Ergebnissen anderer Arbeiter vergleichen können. Diese letzteren Vergleiche können zum Gleichkommen zu anderen Arbeiter oder zum Führerschaften wie auch zu sozialistischem Wettbewerb um Qualitätserhöhung führen, was als wichtigste Aufgabe der SQK gilt.

LITERATUR

1. OBALSKI, J., Statystyczna Kontrola jakosci podczas producji PWT. Warszawa (1955).
2. ODERFELD, J., Zarys statystycznej kontroli jakosci PWN. Warszawa (1954).
3. GRAF, U., MENNING, H. J., Formeln und Tabellen der mathematischen Statistik. Springer Verlag, Berlin, Göttingen (Heidelberg) (1958).

DIE STEUERUNG DER QUALITÄT VON NORMTEILEN MIT MATHEMATISCH-STATISTISCHEN METHODEN

von

H. FINSTERBUSCH

VVB Wälzlager und Normteile
Karl-Marx-Stadt, Deutsche Demokratische Republik

Mit den vorliegenden Zeilen möchten wir einen kurzen Überblick über die Problematik bei der Steuerung der Qualität von Normteilen geben. Obwohl die Norm- oder Standardteileindustrie ihrem Charakter der Massenfertigung nach zu einem der prädestinierten Zweige bei der Anwendung mathematisch-statistischer Methoden gehört, sind doch in der Praxis viele Besonderheiten zu beachten, die eine schematische Übertragung aus anderen Industriezweigen nicht ohne weiteres zulassen.

Der mengenmäßig hohe Ausstoß — vor allem von kaltgeschlagenen Erzeugnissen — zwingt unbedingt zu einer operativen Fertigungssteuerung mittels Kontrollkarten, da selbst eine gut besetzte Endkontrolle unter der Voraussetzung eines zügigen Teiledurchlaufes nicht den geringsten Einfluß mehr auf die Fertigung nehmen kann.

Die fast ausschließliche Verwendung von Lehren zur Prüfung der Maßhaltigkeit von Normteilen bedingt den Einsatz von Kontrollkarten für die Attributprüfung; dabei verwenden wir die Kreuzkarte.

Ist für einen Fertigungsabschnitt die Einführung von Kontrollkarten vorgesehen, erfolgt ein gründlicher Maschinentest, der Auskunft über das qualitative Leistungsvermögen der Maschinen, Anlagen und Werkzeuge gibt. Ohne diese vorherige Überprüfung der Maschinenfunktionsgenauigkeit wird keine Kontrollkarte eingerichtet. In jedem Falle muß gewährleistet sein, daß die Maschine in der Lage ist, die Toleranzen einzuhalten.

Bei der Festlegung der zu überwachenden Qualitätsmerkmale, der Höhe der zu prüfenden Teile (bei Anwendung der Kreuzkarte 5–10%) und des

TABELLE I

Teil	Abmessung		Qualitätsmerkmal
Fehlerart	hervorgerufen durch	abzustellen durch	Fehler-häufigkeit

Prüfturnus stellt sich aber bald heraus, daß der Betrieb mangels genügender Laufkontrolleure nicht in der Lage ist, dieses Programm zu realisieren.

Die ständig steigenden Anforderungen an die Qualität, vor allem an ein gleichbleibendes Qualitätsniveau der Erzeugnisse, stellen aber den normteileproduzierenden Betrieb vor die Aufgabe einer kontinuierlichen Beeinflussung des Fertigungsprozesses. Damit mußte nach Möglichkeiten gesucht werden, den Prüfaufwand durch Verringerung der Stichprobengröße wesentlich zu reduzieren. Die Voraussetzungen dazu bilden gründliche Fehler-Ursachen-Analysen. Diese werden für jedes zu überwachende Qualitätsmerkmal durchgeführt, um festzustellen, ob dieses Merkmal im wesentlichen zufälligen oder systematischen Fehlerursachen unterliegt (Tabelle I).

In Verbindung mit den Ergebnissen der Maschinenfunktionsgenauigkeitsaufnahme, der Gegenüberstellung des natürlichen Streubereiches mit der Zeichnungstoleranz, Trend-Untersuchungen usw., ist es möglich, den praktischen Gegebenheiten angepaßte Stichprobengrößen festzulegen.

Es wird so verfahren, daß bei einem Prüfturnus von einer Stunde 5 bis 10 Teile aus der laufenden Fertigung (je nach dem wie die Ergebnisse aus oben angeführten Untersuchungen ausgefallen sind) zur Überwachung der Qualitätsmerkmale entnommen werden, für die aus den Analysen folgendes bekannt ist:

1. Der natürliche Streubereich ($6\,\sigma$) ist sehr klein im Verhältnis zur Zeichnungstoleranz.

Unter diesen Voraussetzungen kann selbst mit dieser minimalen Stichprobe und unter Verwendung von Lehren rechtzeitig die Überschreitung der Toleranzgrenzen bemerkt werden, da durch die geringe Standardabweichung der Normalverteilung der Ausschußprozentsatz groß genug ist.

Diese Erscheinung tritt bei dem überwiegenden Teil der an kaltgeschlagenen Erzeugnissen zu überwachenden Qualitätsmerkmalen auf, aber auch beim Zerspannungsvorgang (z. B. beim automatischen Gewindeschneiden mit Schneideisen).

2. Ohne nennenswerten vorhandenen Trend tritt plötzlich ein Ereignis ein, das jedes oder entsprechend dem Verfahren jedes i-te Teil fehlerhaft werden läßt.

Das geschieht z. B. durch Bruch einer Matrize infolge Abstumpfung, Versacken von Spänen oder das Ausbrechen von Teilen eines Werkzeuges.

Merkmale, für die das unter diesen beiden Punkten Genannte nicht zu erwarten ist, müssen nach wie vor durch die vollen Stichprobengrößen überwacht werden.

Da mit der kleinen Stichprobe auch alle systematischen Fehler wie Einstellfehler, Fehler in der Werkzeug-oder Materialabmessung usw. mit erfaßt werden, ist es bei kaltgeschlagenen Erzeugnissen möglich, den Prüfaufwand auf 1/10 bis 1/20, in der zerspanenden Fertigung durchschnittlich auf die Hälfte zu reduzieren.

Gleichlaufend mit der verstärkten Anwendung der operativen Kontrolle und dem Einsatz von Kreuzkarten kommt die Forderung nach einem System der Qualitätsbeurteilung während der Fertigung unter Auswertung der Kreuzkarten. Damit soll die Endkontrolle zu Gunsten der laufenden Kontrolle weitestgehend reduziert werden.

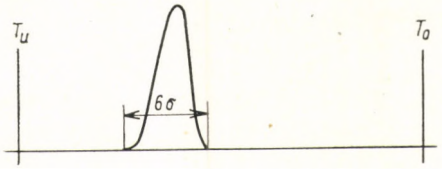


Abb. 1

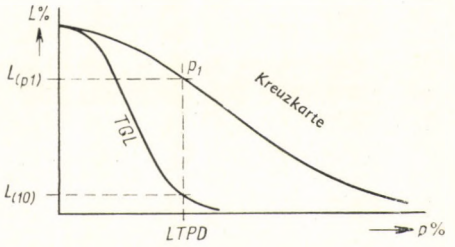


Abb. 2

Schrauben, Muttern und ähnliche Gewinde- und Formteile sowie Niete unterliegen in der DDR technischen Lieferbedingungen, die mathematische Stichprobenpläne enthalten. Damit sind eindeutig die Verhältnisse zwischen Hersteller und Abnehmer festgelegt.

Nun stellt die Kreuzkarte durch die Festlegung der Stichprobengröße n und der Zahl der zulässigen Fehler c ebenfalls einen Stichprobenplan dar, der aber bedingt durch die kleine Stichprobe eine wesentlich flacher verlaufende Operationscharakteristik als der zur Abnahme der Erzeugnisse vorgeschriebene Stichprobenplan besitzt.

Um den Verbraucher vor der Annahme qualitativ minderwertiger Lieferungen zu schützen, ist also nötig, durch geeignete Verfahren die Operationscharakteristik der Kreuzkarte oder wenigstens den Verbrauchersicherheitspunkt LTPD der Operationscharakteristik des vorgeschriebenen Stichprobenplanes anzunähern. Um das zu erreichen, läßt sich bei einer kontinuierlich unter den gleichen Bedingungen laufenden Fertigung das Multiplikationsgesetz der Wahrscheinlichkeitsrechnung ausnutzen.

Es wird verlangt, daß sowohl die erste als auch die zweite bis i-te Stichprobenprüfung nach der Kreuzkarte zur Annahme geführt hat. Ist das der Fall, hat die Prüfung die verlangte Schärfe erreicht. Das bedeutet, daß die Erzeugnisse ohne Endkontrolle das Werk verlassen können. Man bezeichnet diese Methode als Reihenprüfung.

Die Bestimmung des Faktors i
(i ist der Exponent der Gleichung siehe Abb. 2):

$$L_{(10)} = L^i_{(P1)} \,.$$

Wird, wie oft in der Praxis, die Annahmewahrscheinlichkeit mit Hilfe der tabellierten Poissonverteilung berechnet, wird

$$L_{(10)} = \left(\sum_{z=0}^{i} \frac{(np)^z}{z!} \, \mathrm{e}^{-np} \right)^i$$

n und c sind die Parameter der Kreuzkarte. p ist das LTPD des vorgeschriebenen Stichprobenplanes.

Nach i aufgelöst ergibt sich

$$i = \frac{\log L_{(10)}}{\log L_{(P1)}}$$

Mit i erhält man die Anzahl von Teillosen, die nacheinander (von Stunde zu Stunde) nach den Annahmebedingungen der Kreuzkarte zur Annahme geführt haben müssen, um ohne Endkontrolle als »gut« deklariert werden zu können.

Es ist dadurch möglich, die Qualitätsbeurteilung unmittelbar in der Fertigung vom Laufkontrolleur vornehmen zu lassen. Voraussetzungen sind Qualifikation und hohes Qualitäts- und Verantwortungsbewustsein der Laufkontrolleure.

LOCALLY MOST POWERFUL RANK TESTS OF INDEPENDENCE

by

J. HÁJEK

Charles University
Prague, Czechoslovakia

1. THEORY

Let $(X_1, Y_1), \ldots, (X_N, Y_N)$ denote a random sample from a bivariate distribution. Under hypothesis, say H, we shall assume that the random variables X_i and Y_i are independent and distributed according to same unknown densities $f(x)$ and $g(x)$, respectively. Thus H is composite and may be regarded as a family of $2N$-dimensional densities

$$p(x_1, y_1, \ldots, x_N, y_N) = \prod_{i=1}^{N} f(x_i)\, g(y_i)\,,$$

where f and g both run through the set of all one-dimensional densities. The hypothesis H will be tested against a simple alternative

$$q_\Delta(x_1, y_1, \ldots, x_N, y_N) = \prod_{i=1}^{N} h_\Delta(x_i, y_i)\,,$$

where

$$h_\Delta(x, y) = \int_{-\infty}^{\infty} f_0(x - \Delta z)\, g_0(x - \Delta z)\, dM(z)$$

with f_1 and g_0 being some particular densities, and $M(z)$ being a distribution function with positive and finite variance

$$0 < \sigma^2 = \int z^2 dM(z) - \Big(\int z dM(z)\Big)^2 < \infty.$$

Thus, under the alternative, we assume that

$$X_i = X_i^0 + \Delta Z_i\,,$$
$$Y_i = Y_i^0 + \Delta Z_i\,,$$

where X_i^0, Y_i^0, Z_i are independent variables with respective distributions determined by f_0, g_0 and $M(z)$, and Δ denotes a real number.

Since the hypothesis is very wide (non-parametric) we restrict ourselves to rank tests. Moreover, since it is difficult to find explicitly the most powerful rank test, we shall seek the locally most powerful rank test. Recall that a rank test is called locally most powerful for testing H against q_Δ, if it is most powerful for all Δ in some neighbourhood of 0. Further recall that a

test is called a rank test for testing H, if its critical region (or its critical function) is determined by $(R_1, \ldots R_N, Q_1, \ldots, Q_N)$, where R_i denotes the rank of X_i within the sample $(X_1, \ldots, X_N)$ and Q_i denotes the rank of Y_i within the sample $(Y_1, \ldots, Y_N)$. Denoting the ith order statistic for $(X_1, \ldots, X_N)$ by $X^{(i)}$, and the ith order statistic in $(Y_1, \ldots, Y_N)$ by $Y^{(i)}$, we have

$$X_i = X^{(R_i)}, \; Y_i = Y^{(Q_i)}, \; 1 \leqq i \leqq N.$$

Now we must introduce some regularity conditions on f_0 and g_0. We shall assume that f_0 and g_0 are both absolutely continuous on finite intervals, i.e. that there exist functions f_0' and g_0' such that

$$f(x) - f(x_0) = \int_{x_0}^{x} f_0'(t)\, dt$$

and

$$g(x) - g(x_0) = \int_{x_0}^{x} g_0'(t)\, dt\,.$$

Furthermore, we shall suppose that

$$\int_{-\infty}^{\infty} | f_0'(t) |\, dt < \infty \tag{1}$$

and

$$\int_{-\infty}^{\infty} | g_0'(t) |\, dt < \infty \tag{2}$$

and that f_0' and g_0' both are *continuous almost everywhere.*

Finally, introduce the following quantities (called *scores)*:

$$a_N(i, f_0) = E_0\{f_0'(X^{(i)})/f_0(X^{(i)})\}, \; 1 \leqq i \leqq N,$$

where E_0 refers to a sample $X_1, \ldots, X_N$ from a distribution with density f_0.

THEOREM. *Under the above assumptions and notations, the test with the critical region*

$$\sum_{i=1}^{N} a_N(R_i, f_0)\, a_N(Q_i, g_0) \geqq k \tag{3}$$

is the locally most powerful rank test for H against q_Δ on level $\alpha = \alpha\,(k)$, with α equaling the probability of (3) *under H.*

PROOF. Let $r = (r_1, \ldots, r_N)$, $r' = (r_1', \ldots, r_N')$ be some permutations of $(1, \ldots, N)$, and let us put $R = (R_1, \ldots, R_N)$, $Q = (Q_1, \ldots, Q_N)$. Further, let us denote by P_Δ the probability distribution corresponding to the density q_Δ, $dP_\Delta = q_\Delta\, dx$. We shall show that

$$\lim_{\Delta \to 0} \frac{1}{\Delta^2} [(N!)^2 P_\Delta(R = r, Q = r') - 1] = \sigma^2 \sum_{i=1}^{N} a_N(r_i, f_0)\, a_N(r_i', f_0)\,,$$

from which our assertion follows immediately. Actually, then each pair (r, r') such that $\Sigma\, a_N(r_i, f_0)\, a_N(r'_i, g_0) \geq k$ is more probable under P_Δ than each pair (s, s') such that $\Sigma\, a_N(s_i, f_0).\ a_N(s'_i, g_0) < k$, for sufficiently small Δ.

Introduce densities

$$f_\Delta(x) = \int f(x - \Delta z)\, dM(z)$$

$$g_\Delta(x) = \int g(x - \Delta z)\, dM(z)$$

and note that

$$(N!)^2 \int \dots \int_{\substack{R=r \\ Q=r'}} \prod_{i=1}^{N} f_\Delta(x_i)\, g_\Delta(y_i)\, dx_1 \dots dx_N dy_1 \dots dy_N = 1.$$

Consequently,

$$\frac{1}{\Delta^2}[(N!)^2 P_\Delta(R = r, Q = r') - 1] =$$

$$= (N!)^2 \int \dots \int_{\substack{R=r \\ Q=r'}} \Delta^{-2} [\prod_{i=1}^{N} h_\Delta(x_i, y_i) - \prod_{i=1}^{N} f_\Delta(x_i)\, g_\Delta(y_i)] \cdot$$

$$\cdot dx_1 \dots dx_N dy_1 \dots dy_N = \sum_{k=1}^{N} (N!)^2 \int \dots \int_{\substack{R=r \\ Q=r'}} \left[\frac{h_\Delta(x_k, u_k) - f_\Delta(x_k)\, g_\Delta(y_k)}{\Delta^2}\right] \cdot$$

$$\cdot [\prod_{i=k+1}^{N} f_\Delta(x_i) g_\Delta(x_i) \prod_{j=1}^{k-1} h_\Delta(x_j, y_j)]\, dx_1 \dots dx_N dy_1 \dots dy_N . \tag{4}$$

Now, obviously,

$$\lim_{\Delta \to 0} f_\Delta(x)\, g_\Delta(y) = f(x) g(y)$$

$$\lim_{\Delta \to 0} h_\Delta(x, y) = f(x) g(y)$$

in every point (x, y). If we further show that

$$\lim_{\Delta \to 0} \frac{h_\Delta(x, y) - f_\Delta(x) g_\Delta(y)}{\Delta^2} = \sigma^2 f'(x) g'(y) \tag{5}$$

almost everywhere, and that the limit may be carried out under the integral sign, (4) will entail

$$\lim_{\Delta \to 0} \frac{1}{\Delta^2}[(N!)^2 P_\Delta(R = r, Q = r') - 1] =$$

$$= \sum_{k=1}^{N} (N!)^2 \int \dots \int_{\substack{R=r \\ Q=r'}} \sigma^2 \frac{f'_0(x_k)}{f_0(x_k)} \frac{g'_0(y_k)}{g_0(y_k)} \prod_{i=1}^{N} f_0(x_i) g_0(y_i) \cdot dx_1 \dots dx_N dy_1 \dots dy_N =$$

$$= \sigma^2 \sum_{k=1}^{N} \left[N! \int_{R=r} \dots \int \frac{f_0'(x_k)}{f_0(x_k)} \prod_{i=1}^{N} f_0(x_i)\, dx_1 \dots dx_N \right] \cdot$$

$$\cdot \left[N! \int_{Q=r'} \dots \int \frac{g_0'(y_k)}{g_0(y_k)} \prod_{i=1}^{N} g_0(y_i)\, dy_1 \dots dy_N \right] =$$

$$= \sigma^2 \sum_{k=1}^{N} E\left\{\frac{f_0'(X^{(r_k)})}{f_0(X^{(r_k)})}\right\} E\left\{\frac{g_0'(Y^{(r_k')})}{f_0(Y^{(r_k')})}\right\} = \sigma^2 \sum_{k=1}^{N} a_N(r_k, f_0)\, a_N(r_k', g_0)\,.$$

In the above relations we have used the fact that $f(x) = 0$ implies $f'(x) = 0$ in every point, where the derivative exists.

Thus it remains to show that (5) holds almost everywhere and that passing to the limit under the integral sign is justified.

We begin with the proof of (5). We can write

$$h_\Delta(x, y) - f_\Delta(x)\, g_\Delta(y) = \tag{6}$$

$$= \int\int_{z<z'} [f_0(x - \Delta z) - f_0(x - \Delta z')][g_0(y - \Delta z) - g_0(y - \Delta z')]\, dM(z)\, dM(z')\,.$$

If x is a continuity point of f_0', and y is a continuity point of g_0', then

$$[f_0(x - \Delta z) - f_0(x - \Delta z')]\,[g_0(y - \Delta z) - g_0(y - \Delta z')] = \tag{7}$$

$$= \Delta^2 (z - z')^2\, [f_0'(x)\, g_0'(y) + \varepsilon]\,,$$

where $\varepsilon \to 0$ for $\max\{|\Delta z|, |\Delta z'|\} \to 0$.

Furthermore, in view of (1) and (2), the integrand in (6) is uniformly bounded, and

$$\int\int_{\substack{\max\{|\Delta z|, |\Delta z'|\} > \delta \\ z<z'}} dM(z)\, dM(z') \leq (\Delta/\delta)^2 \int_{|z| > \delta/\Delta} z^2 dM(z)\,. \tag{8}$$

Now, (5) almost everywhere obviously follows from (6), (7) and (8), recalling that f_0' and g_0' are continuous almost everywhere and that $\int z^2 dM(z) < \infty$.

In conclusion we show that

$$\limsup_{\Delta \to 0} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Delta^{-2} \,|\, h_\Delta(x_k, y_k) - f_\Delta(x_k)\, g_\Delta(y_k) \,|\, \cdot \tag{9}$$

$$\cdot \prod_{1=k+1}^{N} f_\Delta(x_i)\, g_\Delta(x_i) \prod_{j=1}^{k-1} h_\Delta(x_j, y_j)\, dx_1 \dots dx_N dy_1 \dots dy_N \leqq$$

$$\leqq \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \sigma^2 \,|\, f_0'(x_k) g_0'(y_k) \,|\, \prod_{i \neq k}^{N} f_0(x_i)\, g_0(x_i)\, dx_1 \dots dx_N dy_1 \dots dy_N$$

holds. It may be shown that (9) is a sufficient condition for passing to the limit under the integral sign, with the integral being extended over any

Borel subset, and a fortiori over the subset $\{R = r, Q = r'\}$. First observe that

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Delta^{-2} \, | \, h_\Delta(x_k, y_k) - f_\Delta(x_k) \, g_\Delta(y_k) \, | \cdot$$

$$\cdot \Big[\prod_{i=k+1}^{N} f_\Delta(x_i) \, g_\Delta(x_i) \prod_{j=1}^{k-1} h_\Delta(x_j, y_j) \Big] \, dx_1 \dots dx_N dy_1 \dots dy_N = \tag{10}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Delta^{-2} \, | \, h_\Delta(x_k, y_k) - f_\Delta(x_k) \, g_\Delta(y_k) \, | \, dx_k dy_k$$

and, similarly,

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \sigma^2 \, | \, f_0'(x_k) \, g_0'(y_k) \, | \prod_{i=k}^{N} f_0(x_i) \, g_0(y_i) \, dx_1 \dots dx_N dy_1 \dots dy_N =$$

$$= \sigma^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | \, f_0'(x_k) \, g_0'(y_k) \, | \, dx_k dy_k \, . \tag{11}$$

Further, note that (6) entails

$$| \, h_\Delta(x_k, y_k) - f_\Delta(x_k) \, g_\Delta(y_k) \, | \leqq$$

$$\leqq \int_{z<z'} \int | \, f_0(x - \Delta z) - f_0(x - \Delta z') \, | \, | \, g_0(y - \Delta z) - g_0(y - \Delta z') \, | \, dM(z) \, dM(z') \leqq$$

$$\leqq \int_{z<z'} \int \int_{\Delta z}^{\Delta z'} \int_{\Delta z}^{\Delta z'} | \, f_0'(x - t) \, g_0'(y - s) \, | \, dt ds \, dM(z) \, dM(z') \, .$$

Consequently, interchanging the integration order (which is justified with non-negative integrands), we obtain

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | \, h_\Delta(x_k, y_k) - f_\Delta(x_k) \, g_\Delta(y_k) \, | \, dx_k dy_k \leqq$$

$$\leqq \int_{z<z'} \int \int_{\Delta z}^{\Delta z'} \int_{\Delta z}^{\Delta z'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | \, f_0'(x_k - t) \, g_0'(y_k - t) \, | \, dx_k dy_k \, dt ds dM(z) \, dM(z') =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | \, f_0'(x_k) \, f_0'(y_k) \, | \, dx_k dy_k \int_{z<z'} \int \int_{\Delta z}^{\Delta z'} \int_{\Delta z}^{\Delta z'} dt ds dM(z) \, dM(z') =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | \, f_0'(x_k) \, g_0'(y_k) \, | \, dx_k dy_k \int_{z<z'} \int \Delta^2 (z - z')^2 \, dM(z) \, dM(z') =$$

$$= \Delta^2 \sigma^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | \, f_0'(x_k) \, g_0'(y_k) \, | \, dx_k dy_k \, . \tag{12}$$

Now (9) follows from (10), (11) and (12). Q.E.D.

2. EXAMPLES

If f_0 and g_0 both are logistic densities,

$$f_0 = g_0 = e^{-x}(1 + e^{-x})^{-2},$$

then

$$a_N(i, f_0) = b_N(i, g_0) = \frac{2i}{N+1} - 1$$

and the locally most powerful test coincides with one based on the Spearman's statistic (see [3])

$$\varrho = \frac{12}{N^3 - N} \sum_{i=1}^{N} \left(R_i - \frac{N+1}{2}\right)\left(Q_i - \frac{N+1}{2}\right).$$

If f_0 and g_0 both are normal densities, then the locally most powerful test is based on

$$S = \sum_{i=1}^{N} a_N(R_i)\, a_N(Q_i), \tag{13}$$

where

$$a_N(i) = EZ_N^{(i)}$$

with $Z_N^{(i)}$ being the ith order statistic in a normal sample of size N. Thus S may be called a *rank correlation coefficient with normal scores.*

Very often approximate scores are preferred to exact ones. We put, for example,

$$a_N(i, f_0) \doteq \varphi\left(\frac{i}{N+1}, f_0\right), \tag{14}$$

where

$$\varphi(u, f_0) = -\frac{f'\big(F^{-1}(u)\big)}{f\big(F^{-1}(u)\big)}.$$

Thus, for f_0 and g_0 both normal, (13) may be used with

$$a_N(i) = \Phi^{-1}\left(\frac{i}{N+1}\right)$$

with Φ^{-1} denoting the inverse of the standardized normal distribution function. In this case we could speak about a rank correlation coefficient of van der Waerden type.

Finally, if

$$f_0(x) = g_0(x) = \frac{1}{2} e^{-|x|},$$

then (14) yields

$$a_N(i, f_0) = b_N(i, g_0) \doteq \operatorname{sign}\left(i - \frac{N+1}{2}\right)$$

and the test comes to be the *quadrant* test(see [1]).

3. MOMENTS AND DISTRIBUTION

Consider a general statistic

$$S = \sum_{i=1}^{N} a(R_i)\, b(Q_i)\,.$$

Then under hypothesis,

$$E\,S = N\,\bar{a}\,\bar{b}$$

and

$$\operatorname{var} S = (N-1)^{-1}\,\sigma_a^2\,\sigma_b^2\,,$$

where

$$\bar{a} = \frac{1}{N}\sum_{i=1}^{N} a(i)$$

$$\sigma_a^2 = (N-1)^{-1}\sum_{i=1}^{N}[a(i)-\bar{a}]^2$$

and $\bar{b}$ and σ_b^2 are defined similarly.

If

$$a(i) = a_N(i, f_0)$$

$$b(i) = b_N(i, g_0)$$

where

$$\int_{-\infty}^{\infty}[f_0'(x)/f_0(x)]^2 f_0(x)\,dx < \infty$$

$$\int_{-\infty}^{\infty}[g_0'(x)/g_0(x)]^2 g_0(x)\,dx < \infty\,,$$

then S is asymptotically normal (ES, var S) for $N \to \infty$. This may be derived from the results of [2].

REFERENCES

1. ELANDT, REGINA C., Exact and Approximate Power of the Non-parametric Test of Tendency. *Ann. Math. Stat.* **33,** 471–481 (1962).
2. HÁJEK, J., Some Extensions of the Wald–Wolfowitz–Noether Theorem. *Ann. Math. Stat.* **32,** 506–523 (1961).
3. KENDALL, M. G., Rank Correlation Method. London, Charles Griffin and Co. 2nd ed. (1955).

SOME NEW RESULTS IN THE THEORY OF RANK TESTS

by

J. HÁJEK

Charles University
Prague, Czechoslovakia

1. LOCALLY MOST POWERFUL RANK TESTS

Let H_2 denote the family of $2N$-dimensional densities

$$p = \prod_{i=1}^{N} f(x_i)\, g(y_i)$$

where f and g may be arbitrary one-dimensional densities. Further, let $M(z)$ be a distribution function such that $\int z dM = 0$, $0 < \int z^2\, dM < \infty$, and put

$$h_\Delta(x,y) = \int_{-\infty}^{\infty} f_0(x - \Delta z)\, g_0(x - \Delta z)\, dM$$

and

$$q = \prod_{i=1}^{N} h_\Delta(x_i, y_i).$$

Assume that f_0 and g_0 are absolutely continuous on every finite interval, and that

$$\int_{-\infty}^{\infty} |f_0'(x)|\, dx < \infty\,, \quad \int_{-\infty}^{\infty} |g_0'(x)|\, dx < \infty\,.$$

Then the locally ($\Delta \to 0$) most powerful rank test for H_2 against q is based on the statistic

$$S = \sum_{i=1}^{N} a(R_i)\, b(Q_i)\,,$$

where R_i is the rank of X_i in the ordered x-sample $X^{(1)} < \ldots < X^{(N)}$ and Q_i is the rank of Y_i in $Y^{(1)} < \ldots < Y^{(N)}$,

$$a(i) = E\left\{-\frac{f_0'(X^{(i)})}{f_0(X^{(i)})}\right\}$$

$$b(i) = E\left\{-\frac{g_0'(Y^{(i)})}{g_0(Y^{(i)})}\right\}.$$

2. EXTENSION OF THE KOLMOGOROV–SMIRNOV TEST, AND OF THE RELATED TESTS, TO REGRESSION ALTERNATIVES

Let $(D_1, \ldots, D_N)$ be the inverse permutation to the permutation $(R_1, \ldots, R_N)$, $R_{D_i} = D_{R_i} = i$, $1 \leqq i \leqq N$. Let $c_1, \ldots, c_N$ be real numbers, and put $\bar{c} = N^{-1} \Sigma c_i$. We shall call the numbers R_i ranks, and the numbers D_i antiranks. Obviously $X^{(i)} = X_{D_i}$. Now put

$$S_c = \frac{\max\limits_{1 \leqq k \leqq N} (k\bar{c} - c_{D_1} - \ldots - c_{D_k})}{\left[\sum\limits_{i=1}^{N} (c_i - \bar{c})^2\right]^{\frac{1}{2}}}.$$

For the special case, in which $c_i = 1$, $1 \leqq i \leqq m$, and $c_i = 0$, $m \leqq i \leqq m + n = N$, this statistic coincides with the one-sided Kolmogorov–Smirnov statistic for comparing two samples. The two-sided statistic, as well as the Rényi statistic and the Cramér–Mises statistic may be defined similarly. The statistic S_c is appropriate for testing randomness against the alternative

$$q = \prod_{i=1}^{N} f(x_i - \Delta c_i).$$

THEOREM. *If*

$$\frac{\sum\limits_{i=1}^{N} (c_i - \bar{c})^2}{\max\limits_{1 \leqq i \leqq N} (c_i - \bar{c})^2} \to \infty,$$

then

$$P(S_c < \lambda) \to 1 - e^{-2\lambda^2}.$$

Similar assertions hold for other above-mentioned statistics. For the two sample cases when $c_i = 1$, $1 \leqq i \leqq m$, and $c_i = 0$, $m < i \leqq m + n$, (1) is equivalent to

$$m \to \infty, \quad n \to \infty$$

whichever the ratio m/n.

3. LOCAL ASYMPTOTIC EFFICIENCY

Let $H_{0,N}$ denote the family of densities

$$p = \prod_{1}^{N} f(x_i),$$

where f may be arbitrary one-dimensional density. Now consider two parallel testing situations. In the first situation we test $H_{0,m+n}$ against

$$q = \prod_{i=1}^{m} f_0(x_i - \Delta) \prod_{i=m+1}^{m+n} f_0(x_i),$$

where

$$\Delta^2 \frac{m+n}{n} = b > 0$$

by an asymptotically most powerful test. In the second situation we test $H_{0,m'+n'}$ against

$$q = \prod_{i=1}^{m'} f_0(x_i - \Delta) \prod_{i=m'+1}^{m'+n'} f_0(x_i),$$

where

$$m = [em']$$

$$n = [en']$$

by the Kolmogorov–Smirnov test. We let $m \to \infty$, $n \to \infty$ and ask for what e the powers of the two tests are asymptotically equal. The solution will depend on the level α and b from (2), $e = e(\alpha, b)$, and will be complicated. The limit

$$e(\alpha) = \lim_{b \to 0} e(\alpha, b)$$

is simpler, and will be called the *local asymptotic efficiency* of the Kolmogorov–Smirnov test.

THEOREM. *It holds*

$$e(\alpha) = [I(f_0)]^{-1} \exp\ (k^2_{1-\alpha}) 4\pi\alpha^2 \log \alpha^{-1} \cdot$$

$$\cdot \int_0^1 \varphi(t, f_0) \left[2\Phi\left(\frac{\lambda_{1-\alpha}(2t-1)}{t(1-t)} \right) - 1 \right] dt$$

where

$$I(f_0) = \int_{-\infty}^{\infty} [f_0'(x)/f_0(x)]^2 f_0(x)\, dx$$

Φ is the standardized normal distribution function,

$$\Phi(k_{1-\alpha}) = 1 - \alpha, \quad \exp(-2\lambda^2_{1-\alpha}) = \alpha,$$

and

$$\varphi(t, f_0) = - \frac{f_0'\big(F^{-1}(t)\big)}{f_0\big(F^{-1}(t)\big)}.$$

We have

$$\sup_{f_0} e\,(0{\cdot}05) = 0{\cdot}852$$

$$\sup_{f_0} e\ (0{\cdot}01) = 0{\cdot}864$$

$$\limsup_{\alpha \to 0} \sup_{f_0} e\ (\alpha) = 1$$

where the supremum is taken over all f_0 with $I(f_0) < \infty$.

REFERENCE

1. HÁJEK, J., ŠIDÁK, Z., Theory of Rank Tests, Academia, Praha, 1967.

ALLGEMEINE GRENZWERTSÄTZE DER WAHRSCHEINLICHKEITSRECHNUNG

von

H. HEYER

Institut für Versicherungsmathematik und Mathematische Statistik der Universität Hamburg, Deutsche Bundesrepublik

1. KURZE ZUSAMMENFASSUNG DER ERGEBNISSE

Der Begriff der Verteilungsfunktion einer Zufallsvariablen erweist sich als schwerfällig zu handhaben, wenn man als Bildmenge der Zufallsvariablen anstelle der reellen Zahlen den $\mathbf{R}^N$ für $N \geq 2$ zuläßt, und verliert seinen Sinn gänzlich, wenn man sich für Bildmengen mit allgemeineren Strukturen interessiert, etwa für solche ohne Ordnungsstruktur. Deshalb ist man inzwischen dazu übergegangen, anstelle der Verteilungsfunktion die Verteilung selbst als Maß auf der Bildmenge der Zufallsvariablen in den Vordergrund der Theorie zu stellen. Dabei zeigt sich, daß man ohne große zusätzliche Mühen sofort beim Begriff der Wahrscheinlichkeitsverteilung auf lokalkompakten topologischen Räumen angelangt ist. Die Verteilungskonvergenz in der Menge aller Verteilungsfunktionen wird ersetzt durch die allgemeinere vage Konvergenz in der Menge aller W-Verteilungen.

Es ergeben sich innerhalb der so erweiterten Theorie die folgenden Grenzwertsätze, die für die W-Rechnung sowohl als auch für die Statistik ihre eigene Bedeutung haben: Das Studium der wichtigsten Eigenschaften quadrierbaren Mengen erlaubt den Beweis gleichwertiger Aussagen für die vage Konvergenz beliebiger positiver Maße. Standardbeweise liefern Zusammenhänge zwischen vager und Bernoulli-Konvergenz (B-Konvergenz). Als entscheidendes Hilfsmittel für das Studium unteilbarer W-Verteilungen läßt sich ein Satz beweisen, der die B-Konvergenz durch die gleichmäßige Konvergenz auf Teilmengen der Menge aller stetigen beschränkten Funktionen charakterisiert, die im Sinne der gleichmäßigen Konvergenz kompakt sind. Bei gleichmäßiger Dominanz durch ein Borelmaß folgt aus der vagen Konvergenz die Konvergenz auf den Borelschen Mengen. Ferner zeigt sich, daß Folgen gleichmäßig absolut stetiger W-Verteilungen bzgl. desselben dominanten Maßes absolut stetige Limiten besitzen. Für den Fall eines lokalkompakten Raumes mit Gruppenstruktur (einer lokalkompakten Gruppe) lassen sich weitere nützliche Grenzwertsätze beweisen: Aus der stochastischen Konvergenz einer Folge gruppenwertiger Zufallsvariablen (G-Zufallsvariablen) folgt die vage Konvergenz der Folge ihrer Verteilungen. Die stochastische Konvergenz gegen die konstante Zufallsvariable e (Einheitselement der Gruppe) ist ferner äquivalent mit der vagen Konvergenz der Verteilungen gegen das Punktmaß im Einheitselement.

Schließlich ergeben sich die Stetigkeit der Faltungsabbildung und ein Satz über Translationskompaktheit in der Menge der Faltungsprodukte zweier W-Verteilungen bzgl. der Topologie der B-Konvergenz.

2. MAßTHEORETISCHE VORBEMERKUNGEN

Der folgenden Darstellung liegt die Terminologie von N. Bourbaki [5] zugrunde. Sei X ein lokalkompakter Raum und seien $\mathscr{K} = \mathscr{K}(X)$ bzw. $\mathscr{C}^b = \mathscr{C}^b(X)$ bzw. $\mathscr{C}^\infty = \mathscr{C}^\infty(X)$ die Vektorräume der reellwertigen stetigen Funktionen mit kompaktem Träger bzw. stetigen beschränkten bzw. stetigen im Unendlichen verschwindenden Funktionen auf X. Mit $\mathscr{K}_+$, $\mathscr{C}^b_+$ und $\mathscr{C}^\infty_+$ werden die Mengen der positiven Funktionen aus $\mathscr{K}$, $\mathscr{C}^b$ und $\mathscr{C}^\infty$ bezeichnet. $\mathscr{K}$, $\mathscr{C}^b$ und $\mathscr{C}^\infty$ seien mit der Normtopologie ausgerüstet. (Radonsches) *Maß* auf X heißt dann jede relativ beschränkte Linearform μ auf $\mathscr{K}$, d.h. jede Linearform μ auf $\mathscr{K}$ mit der Eigenschaft $\sup_{|f| \leq g} |\mu(f)| < \infty$ für alle $g \in \mathscr{K}_+$. *Beschränkt* heißt ein Maß μ auf X, wenn es stetig ist auf $\mathscr{K}$ bezüglich der Topologie der gleichmäßigen Konvergenz.

Als Abkürzung benützen wir $\mathscr{M} = \mathscr{M}(X)$ für die Menge aller Maße, $\mathscr{M}^b = \mathscr{M}^b(X)$ für die Menge aller beschränkten Maße sowie $\mathscr{M}_+ = \mathscr{M}_+(X)$ für die Menge aller positiven Maße auf X. Darüber hinaus bezeichnet $\mathscr{M}^b_+ = \mathscr{M}^b_+(X)$ die Menge aller positiven und beschränkten Maße sowie $\mathscr{M}^1_+ = \mathscr{M}^1_+(X)$ die Menge aller W-Maße auf X.

Es wird $\mathscr{M}^b$ zu einem Banachraum, wenn man die Norm $||\mu|| = \sup_{||f|| \leq 1} |\mu(f)|$ für $f \in \mathscr{K}$ einführt. Notwendig und hinreichend für die Beschränktheit von μ ist die Endlichkeit von $||\mu||$. Aus $\mu \in \mathscr{M}^b_+$ folgt $|\mu(f)| \leq ||\mu|| \, ||f||$ für alle $f \in \mathscr{K}$. Schließlich ist $\mathscr{M}^b_+$ linearer Unterraum von $\mathscr{M}$. Die positiven Maße können nach dem Satz von Riesz als Borelmaße auf den Borelschen Mengen $\mathfrak{B} = \mathfrak{B}(X)$ von X, den Elementen der kleinsten σ-Algebra über den kompakten Teilmengen von X, aufgefaßt werden. *Borelmaße* auf $\mathfrak{B}$ sind also alle Maße, die auf den kompakten Teilmengen von X endlich sind. Für den Fall eines lokalkompakten Raumes mit abzählbarer Basis — und diese Voraussetzung werden wir für alle Zukunft machen — stimmen die Borelschen Mengen mit den Baireschen, den Elementen der kleinsten σ-Algebra über den kompakten G_δ-Mengen in X, überein. Die offenen und abgeschlossenen Teilmengen von X sind Borelsch. Jedes Borelmaß ist durch die Werte auf den kompakten Teilmengen von X bereits eindeutig festgelegt, und es ist *regulär*, d.h.

$$\mu(B) = \inf_{\substack{U \supset B \\ U \text{ offen}}} \mu(U) = \sup_{\substack{K \subset B \\ K \text{ kompakt}}} \mu(K).$$

Insbesondere gibt es zu jedem $\varepsilon > 0$ und jedem $B \in \mathfrak{B}$ eine abgeschlossene Menge $K \subset B$ und eine offene Menge $U \supset B$ mit $\mu(U - K) < \varepsilon$. Für $\mu \in \mathscr{M}^b_+$ kann K kompakt gefunden werden. Zwei Maße $\mu, \nu \in \mathscr{M}_+$ heißen *äquivalent*, wenn sie dieselben lokalen Nullmengen besitzen, d.h. wenn lokal μ-fast überall gilt: $\nu = g\,\mu$ *, wobei g lokal μ-integrierbar und strikt positiv ist. Ist nun X abzählbar im Unendlichen (σ-kompakt), so existiert eine stetige strikt positive Funktion h, so daß das zu μ äquivalente

* Man spricht in diesem Fall von einem Maß ν mit der Dichte g bezüglich μ.

Maß $\nu := h\mu$ beschränkt ist. *Vage Konvergenz* auf $\mathcal{M}$ nennen wir die punktweise Konvergenz in $\mathcal{X}$. Betrachtet man die punktweise Konvergenz in $\mathcal{C}^b$, so spricht man von *Bernoulli-* oder kurz *B-Konvergenz* in $\mathcal{M}$. Die Topologie der vagen Konvergenz auf $\mathcal{M}$ wird durch die Halbnormen $\mu \to \sup_{1\le i\le n} |\mu(f_i)|$ für $(f_i)_{1\le i\le n} \subset \mathcal{X}$ definiert; sie ist lokalkonvex und Hausdorffsch. Besitzt X eine abzählbare Basis, so ist $\mathcal{M}_+$ metrisierbar mit abzählbarer Basis und umgekehrt. Man kann zudem die Metrisierung vollständig wählen. Eine Folge $(\mu_n) \subset \mathcal{M}$ konvergiert vag gegen $\mu \in \mathcal{M}$, in Zeichen $\mu_n \xrightarrow{\text{vag}} \mu$, wenn $\lim_{n\to\infty} \mu_n(f) = \mu(f)$ ist für alle $f \in \mathcal{X}$. Wird $\mathcal{X}$ durch $\mathcal{C}^b$ ersetzt, so vereinbaren wir die Schreibweise $\mu_n \xrightarrow{B} \mu$. Die Abbildung $\mu \to \mu(f)$ ist eine im Sinne der vagen Topologie stetige Linearform auf $\mathcal{M}$. Speziell ist $\mathcal{M}_+$ abgeschlossen in $\mathcal{M}$ bezüglich der vagen Topologie. Definiert man $\mathcal{N} \subset \mathcal{M}$ als *beschränkt*, wenn für alle $f \in \mathcal{X}$ $\sup_{\mu\in\mathcal{N}} |\mu(f)| < \infty$ ist, so erhält man das Ergebnis: Jede vag konvergente Folge aus $\mathcal{M}$ ist beschränkt. In der Tat ist jede beschränkte Teilmenge $\mathcal{N}$ von $\mathcal{M}$ relativ kompakt bezüglich der vagen Topologie (siehe [5], S. 62). Hieraus folgt, daß die Menge $\{\mu \in \mathcal{M}^b;\ \|\mu\| \le a,\ a \in \mathbf{R}^1,\ a > 0\}$ vag kompakt ist. Schließlich ist $\mathcal{M}^1_+$ genau dann vag kompakt, wenn X kompakt ist. ε_x bezeichnet das *Punktmaß* mit der Masse 1 in x ($x \in G$).

Das Ende eines Beweises wird mit $\perp$ angedeutet.

3. VAGE UND BERNOULLI-KONVERGENZ

Über vage und Bernoullikonvergenz in $\mathcal{M}_+$ beweisen wir die drei folgenden Sätze, von denen die ersten beiden für $X = \mathbf{R}^N$ zum Beispiel in [3] stehen. Der dritte Satz stellt eine Verallgemeinerung eines Satzes von R. R. Rao [12] dar, die darin besteht, daß auf die gleichmäßige Beschränktheit der dort auftretenden kompakten Teilmengen $\mathcal{A}$ von $\mathcal{C}^b$ verzichtet wird. Der Beweis zum Satz von R. R. Rao war dem Verfasser zudem unzugänglich.

SATZ 3.1. *Seien* $\mu_n, \mu \in \mathcal{M}_+ (n = 1, 2, \ldots)$. *Es sei ferner* $\mu_n \xrightarrow{\text{vag}} \mu$,

$$\sup_n \mu_n(X) < \infty \text{ für alle } n = 1, 2, \ldots$$

Dann gilt $\lim_{n\to\infty} \mu_n(f) = \mu(f)$ *für alle* $f \in \mathcal{C}^\infty$.

BEWEIS. Wir definieren

$$\alpha := \sup_n \mu_n(X),\ \mu(X) := \sup_{0\le g\le 1} \mu(g)$$

für alle $g \in \mathcal{X}$. Wegen $\mu_n(g) \le \alpha$ folgt

$$\mu(g) = \lim_{n\to\infty} \mu_n(g) \le \alpha$$

und weiter $\mu(X) \leq \alpha$. Seien nun $f \in \mathcal{C}^\infty$ sowie, $\varepsilon > 0$ vorgegeben. Dann gibt es ein $h \in \mathcal{K}$ mit $| f - h | \leq \varepsilon$. Also ergibt sich sowohl $| \mu_n(f) - \mu_n(h)| \leq \leq \varepsilon \alpha$ als auch $| \mu(f) - \mu(h) | \leq \varepsilon \alpha$, woraus

$$| \mu_n(f) - \mu(f) | \leq 2\varepsilon\alpha + \varepsilon$$

folgt für alle $f \in \mathcal{C}^\infty$, $n \geq n_\varepsilon$. ⌟

Bemerkung 3.1. Ohne die Bedingung $\sup_n \mu_n(X) < \infty$ ist der Satz falsch, wie das Beispiel $X = \mathbf{R}^1$, $\mu_n = n\,\varepsilon_n$ und $f(x) = 1$ für $| x | \leq 1$ sowie $| x |^{-1}$ für $| x | \geq 1$ zeigt.

Satz 3. 2. *Eine Folge* (μ_n) *aus* $\mathcal{M}_+$ *konvergiere vag gegen ein Maß* $\mu \in \mathcal{M}_+$ *mit* $\mu(X) < \infty$. *Ferner sei* $\lim_{n\to\infty} \mu_n(X) = \mu(X)$ *erfüllt. Dann konvergiert die Folge* (μ_n) *gegen* μ *auch im Sinne Bernoullis.*

Beweis. O. B. d. A. setzen wir $\mu_n(X) < \infty$ voraus für alle $n = 1, 2, \ldots$. Wir definieren erneut

$$\alpha := \mu(X) = \sup_{0 \leq g \leq 1} \mu(g) \text{ für alle } g \in \mathcal{K}.$$

Es existiert zu jedem $\varepsilon > 0$ ein $g \in \mathcal{K}$ mit $0 \leq g \leq 1$ und $\mu(1 - g) = \alpha - \mu(g) < \varepsilon$. Nach Voraussetzung haben wir

$$\lim_{n\to\infty} \mu_n(X) = \mu(X)$$

und

$$\lim_{n\to\infty} \mu_n(g) = \mu(g) \text{ für alle } g \in \mathcal{K}$$

gegeben. Also ergibt sich $\mu_n(1 - g) < \varepsilon$ für alle $n \geq n_\varepsilon$. Sei nun $f \in \mathcal{C}^b$. Wir können o. B. d. A. $f \geq 0$, $f \leq \beta$, $\beta \in \mathbf{R}^1$, voraussetzen. Aus $f(1 - g) \geq 0$ folgt

$$0 \leq \mu_n\big(f(1 - g)\big) \leq \beta\mu_n(1 - g)$$

für alle $n \geq n_\varepsilon$ und

$$0 \leq \mu\big(f(1 - g)\big) \leq \beta\mu(1 - g) \leq \beta\varepsilon\,.$$

Mithin resultiert

$$\mu_n(fg) \leq \mu_n(f) \leq \beta\varepsilon + \mu_n(fg) \text{ für alle } n \geq n_\varepsilon$$

und analoge Ungleichung für μ anstelle von μ_n. Ist nun $g \in \mathcal{K}$, so folgt

$$fg \in \mathcal{K}\,,$$

und

$$\lim_{n\to\infty} \mu_n(fg) = \mu(fg)$$

impliziert

$$\lim_{n\to\infty} \mu_n(f) = \mu(f) \text{ für alle } f \in \mathcal{C}^b. \text{ ⌟}$$

BEMERKUNG 3. 2. Wegen $\mathscr{X} \subset \mathscr{C}^b$ ergibt sich aus Satz 3.2 eine Charakterisierung der Bernoullikonvergenz beschränkter positiver Maße auf $\mathfrak{B}$ wie folgt: (μ_n) konvergiert im Sinne Bernoullis gegen μ genau dann, wenn (μ_n) vag gegen μ konvergiert und $\lim_{n\to\infty} \mu_n(X) = \mu(X)$ erfüllt ist.

SATZ 3.3. *Seien* $\mu_n, \mu \in \mathscr{M}^1_+ (n = 1, 2, \ldots)$ *mit* $\mu_n \xrightarrow{B} \mu$. *Dann ist diese Konvergenzaussage gleichwertig mit der gleichmäßigen Konvergenz der Folge* $(\mu_n(f))$ *gegen* $\mu(f)$ *für alle* f *aus allen kompakten Teilmengen* $\mathscr{A}$ *von* $\mathscr{C}^b$, *d. h. also:* $\mu_n \xrightarrow{B} \mu \Leftrightarrow \sup_{f\in\mathscr{A}} | \mu_n(f) - \mu(f) | \to 0$ *für alle kompakten Teilmengen* $\mathscr{A}$ *von* $\mathscr{C}^b$.

BEWEIS. 1. Um zu sehen, daß aus der gleichmäßigen Konvergenz auf kompakten Teilmengen $\mathscr{A} \subset \mathscr{C}^b$ der Folge $(\mu_n(f))$ gegen $\mu(f)$ die Bernoullikonvergenz von (μ_n) gegen μ folgt, wähle man einfach $\mathscr{A} = \{f\}, f \in \mathscr{C}^b$.

2. Für die Umkehrung muß gezeigt werden, daß aus der punktweisen Konvergenz auf $\mathscr{C}^b$ die gleichmäßige Konvergenz auf den kompakten Teilmengen von $\mathscr{C}^b$ folgt. Da $\mathscr{C}^b$ vollständig ist, stimmen die Begriffe kompakt und präkompakt überein. Es existieren also zu jedem $\delta > 0$ Elemente $f_1, \ldots, f_n \in \mathscr{A}$ derart, daß es zu jedem $f \in \mathscr{A}$ ein i mit $0 \leq i \leq n$ gibt, so daß $| f - f_i | < \delta$ gilt. Wegen $| \mu(f) | \leq || \mu || \, || f || \leq || f ||$ auf $\mathscr{C}^b$ ist insbesondere jede Teilmenge von $\mathscr{M}^1_+$ gleichgradig stetig. Nimmt man noch die Voraussetzung der Bernoullikonvergenz von (μ_n) gegen μ hinzu, so wird bei

$$\delta = \frac{\varepsilon}{3}$$

$$| \mu_n(f) - \mu(f) | \leq | \mu_n(f) - \mu_n(f_i) | + | \mu_n(f_i) - \mu(f_i) | +$$

$$+ | \mu(f_i) - \mu(f) | < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

für alle $f \in \mathscr{A}$. ⊣

BEMERKUNG 3.3. Ganz analog beweist man die Aussage des Satzes für eine vag konvergente Folge $(\mu_n) \subset \mathscr{M}^b_+$.

4. QUADRIERBARKEIT UND VAGE KONVERGENZ

Um zu äquivalenten Fassungen der vagen Konvergenz in $\mathscr{M}_+$ zu gelangen, geben wir folgende.

DEFINITION 4.1. Sei $\mu \in \mathscr{M}_+$. Eine relativ kompakte Teilmenge A von X heißt *μ-quadrierbar* (oder *μ-Stetigkeitsmenge*), wenn ihr Rand das μ-Maß 0 hat.

Man zeigt leicht, daß eine Menge $A \subset X$ genau dann μ-quadrierbar ist, wenn es zu jedem $\varepsilon > 0$ Funktionen $g, h \in \mathscr{X}$ gibt mit $g \leq \chi_A \leq h$ und $\mu(g - h) < \varepsilon$.

Es ergeben sich nach Bauer [2], (S. 462, 463) die folgenden *Eigenschaften μ-quadrierbarer Mengen*:

1. Mit $A \subset X$ sind auch $\bar{A}$ und $\mathring{A}$ μ-quadrierbar, und es gilt $\mu(A) = \mu(\bar{A}) = \mu(\mathring{A})$.

2. Zu jeder μ-quadrierbaren Menge $A \subset X$ und jeder reellen Zahl $\varepsilon > 0$ existiert eine offene μ-quadrierbare Menge U mit $A \subset U$ und $\mu(U - A) \leq \varepsilon$.

3. Zu jeder kompakten Menge $K \subset X$ und jeder offenen Menge $U \subset X$ mit $K \subset U$ existiert eine μ-quadrierbare Menge $A \subset X$ mit $K \subset A \subset U$.

4. Zu jeder μ-quadrierbaren Menge A und beliebigem $\varepsilon > 0$ existieren eine kompakte μ-quadrierbare Menge $K \subset X$ und eine offene μ-quadrierbare Menge $U \subset X$ mit $K \subset A \subset U$ und $\mu(U - K) \leq \varepsilon$.

Der folgende Satz wird in [1], § 6 sogar für Filter von Maßen in $\mathscr{M}_+$ angekündigt.

SATZ 4.1. *Es sei (μ_n) eine Folge in $\mathscr{M}_+$, $\mu \in \mathscr{M}_+$. Dann sind folgende Aussagen äquivalent:*

(1) *(μ_n) konvergiert vag gegen μ.*

(2) $\lim\limits_{n\to\infty} \mu_n(A) = \mu(A)$ *für alle μ-quadrierbaren, kompakten Mengen $A \subset X$.*

(2′) $\lim\limits_{n\to\infty} \mu_n(A) = \mu(A)$ *für alle μ-quadrierbaren, offenen Mengen $A \subset X$.*

(2″) $\lim\limits_{n\to\infty} \mu_n(A) = \mu(A)$ *für alle μ-quadrierbaren Borelschen Mengen $A \subset X$.*

(3) *Für jede kompakte Menge $K \subset X$ und jede relativ kompakte offene Menge $G \subset X$ gilt:* $\limsup\limits_n \mu_n(K) \leq \mu(K)$ *und* $\liminf\limits_n \mu_n(G) \geq \mu(G)$.

(4) *Sei K kompakte und G relativ kompakte offene Menge in X mit $K \subset G$. Dann existiert zu jedem $\mu_n \in \mathscr{M}_+$ eine kompakte μ-quadrierbare Menge K_n mit $K \subset K_n \subset G$ und eine offene μ-quadrierbare Menge G_n mit $K \subset G_n \subset G$ derart, daß* $\lim\limits_{n\to\infty} \mu_n(K_n) = \mu(K)$ *und* $\lim\limits_{n\to\infty} \mu_n(G_n) = \mu(G)$ *gilt.*

BEWEIS. Wir begnügen uns mit dem Beweis der Äquivalenzen (1) $\Leftrightarrow$ (2″) $\Leftrightarrow$ $\Leftrightarrow$ (3). Zunächst: (1) $\Rightarrow$ (2″): Sei A μ-quadrierbar. Dann existieren zu beliebig vorgegebenem $\varepsilon > 0$ Funktionen $g, h \in \mathscr{K}$ mit $g \leq \chi_A \leq h$ und $\mu(h - g) \leq \varepsilon$, d. h. $\mu_n(g) \leq \mu_n(A) \leq \mu_n(h)$ für alle $n = 1, 2, \ldots$. Nach (1) gilt

$$\lim_{n\to\infty} \mu_n(g) = \mu(g) \text{ und } \lim_{n\to\infty} \mu_n(h) = \mu(h),$$

woraus geschlossen werden kann: $\mu(g) \leq \liminf\limits_n \mu_n(A) \leq \limsup\limits_n \mu_n(A) \leq$ $\leq \mu(h)$. Hieraus folgt $\limsup\limits_n \mu_n(A) - \liminf\limits_n \mu_n(A) \leq \varepsilon$ für alle $\varepsilon > 0$, also existiert $\lim\limits_{n\to\infty} \mu_n(A)$, und es gilt wegen $\mu(g) \leq \mu(A) \leq \mu(h)$ die Behauptung.

Weiter werde (2″) $\Rightarrow$ (3) gezeigt: Sei K kompakt in X. Dann gibt es wegen der Regularität von μ und nach Eigenschaft (3) μ-quadrierbarer Mengen eine μ-quadrierbare Menge $A \in \mathfrak{B}$ mit $K \subset A$ und $\mu(A - K) \leq \varepsilon$. Aus $K \subset A$ folgt $\mu_n(K) \leq \mu_n(A)$ für alle $n = 1, 2, \ldots$. Also hat man

$$\limsup_n \mu_n(K) \leq \limsup_n \mu_n(A) = \lim_{n\to\infty} \mu_n(A) = \mu(A) \leq \mu(K) + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, erhält man $\limsup_n \mu_n(K) \leq \mu(K)$. Auf ähnliche Weise gewinnt man noch die fehlende Ungleichung für relativ kompakte, offene Teilmengen G von X.

Beweis der Implikation (3) $\Rightarrow$ (2″): Sei A μ-quadrierbar und Borelsch. Dann ist $G = \mathring{A}$ offene, relativ kompakte und $K = \bar{A}$ kompakte Menge in X. Man erhält $\limsup_n \mu_n(\bar{A}) \leq \mu(\bar{A})$ und $\liminf_n \mu_n(\mathring{A}) \geq \mu(\mathring{A})$. Wegen $\mathring{A} \subset A \subset \bar{A}$ ist

$$\limsup_n \mu_n(A) \leq \limsup_n \mu_n(\bar{A}) \leq \mu(A) \leq \liminf_n \mu_n(\mathring{A}) \leq \liminf_n \mu_n(A).$$

In jedem Falle gilt zudem: $\liminf_n \mu_n(A) \leq \limsup_n \mu_n(A)$. Also folgt die Behauptung.

Dem Beweis der Implikation (2″) $\Rightarrow$ (1) schicken wir ein Lemma voraus, das eine spezielle Fassung eines Hilfssatzes von Loomis darstellt (S. 171 in [9]).

LEMMA 4.1. Sei $\mu \in \mathscr{M}_+$ und $f \in \mathscr{K}$. Dann sind die Mengen $[f > \xi] = \{x : f(x) > \xi\}$ μ-quadrierbar für alle $\xi \in \mathbf{R}^1$, $\xi > 0$ mit höchstens abzählbar vielen Ausnahmen.

Fortsetzung des Beweises von Satz 4.1. Es sei $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ für alle μ-quadrierbaren Borelschen Mengen $A \subset X$ erfüllt. Zu zeigen ist: $\lim_{n\leftarrow\infty} \mu_n(f) = \mu(f)$ für alle $f \in \mathscr{K}$. O. B. d. A. können wir $f \in \mathscr{K}_+$ voraussetzen. Sei G relativ kompakt, offen, μ -quadrierbar und so gewählt, daß $f(x) = 0$ ist für alle $x \in G^c$. $\varepsilon > 0$ sei beliebig vorgegeben. Dann existiert eine Folge nichtnegativer reeller Zahlen $(\xi_n)_{0 \leq n \leq N}$ mit

$$0 = \xi_0 < \xi_1 < \ldots < \xi_N = \sup_{x \in G} f(x), \Delta_n := \xi_n - \xi_{n-1} \leq \varepsilon$$

und

$$A_n := [f > \xi_n]$$

μ-quadrierbar für alle $n = 1, 2, \ldots, N$ nach Lemma 4.1. Für $n = 0$ setzen wir $A_0 = G$. Dann gilt $\varnothing \subset A_N \subset A_{N-1} \subset \ldots \subset A_0 = G$. Für alle $x \in X$ haben wir

$$\sum_{n=1}^{N} \Delta_n \chi_{A_n}(x) - \varepsilon \leq f(x) \leq \sum_{n=1}^{N} \Delta_n \chi_{A_n}(x) + \varepsilon .$$

Wegen $[f \neq 0] \subset A_0 = G$ gilt weiter

$$\sum_{n=1}^{N} \Delta_n \chi_{A_n} - \varepsilon \chi_{A_0} \leq f \leq \sum_{n=1}^{N} \Delta_n \chi_{A_n} + \varepsilon \chi_{A_0} .$$

Daraus ergibt sich

$$\underline{S}(\mu_m, \varepsilon) := \sum_{n=1}^{N} \Delta_n \mu_m(A_n) - \varepsilon \mu_m(A_0) \leq \mu_m(f) \leq \sum_{n=1}^{N} \Delta_n \mu_m(A_n) + \\ + \varepsilon \mu_m(A_0) =: \bar{S}(\mu_m, \varepsilon)$$

für alle $m = 1, 2, \ldots$. Da $A_0, \ldots, A_n$ μ-quadrierbar und offen sind, erhalten wir aus

$$\lim_{m \to \infty} \mu_m(A_n) = \mu(A_n)$$

für $n = 0, \ldots, N$ die Ungleichung

$$\underline{S}(\mu, \varepsilon) \leq \liminf_m \mu_m(f) \leq \limsup_m \mu_m(f) \leq \bar{S}(\mu, \varepsilon) .$$

Nun ist einerseits

$$\bar{S}(\mu, \varepsilon) - \underline{S}(\mu, \varepsilon) = 2\varepsilon\mu(A_0) ,$$

andererseits

$$\overline{S}(\mu, \varepsilon) \leq \mu(f) \leq \bar{S}(\mu, \varepsilon) .$$

Da A_0 unabhängig von $\varepsilon > 0$ ist, ergibt sich die Behauptung aus

$$\liminf_m \mu_m(f) = \limsup_m \mu_m(f) = \lim_{m \to \infty} \mu_m(f) = \mu(f) . \perp$$

5. GLEICHGRADIGE STETIGKEIT UND VAGE KONVERGENZ

Für den Beweis der folgenden zwei Sätze empfiehlt sich eine Ergänzung zur Definition der absoluten Steigkeit eines Maßes bezüglich eines zweiten.

DEFINITION 5.1. Sei λ ein Borelmaß auf $\mathfrak{B}$, I ein ebeliebige Indexmenge. Eine Familie $(\mu_i)_{i \in I} \subset \mathscr{M}^1_+$ heißt *gleichgradig λ-stetig*, wenn es zu jedem $\varepsilon > 0$ ein $\varepsilon' > 0$ gibt derart, daß aus $\lambda(B) < \varepsilon'$ folgt $\mu_i(B) < \varepsilon$ für alle $B \in \mathfrak{B}$ und alle $i \in I$. Allgemeiner heißt eine Familie $(\mu_i)_{i \in I} \subset \mathscr{M}^1_+$ *$(\varepsilon, \varepsilon')$-gleichgradig λ-stetig*, wenn für reelle Zahlen $\varepsilon, \varepsilon' > 0$ aus $\lambda(B) < \varepsilon'$ folgt: $\mu_i(B) < \varepsilon$ für alle $B \in \mathfrak{B}$ und alle $i \in I$.

Die Spezialisierung der Definition 5.1 auf Folgen in $\mathscr{M}^1_+$ ist evident.

SATZ 5.1. *Seien $\mu, \mu_n \in \mathscr{M}^1_+$ $(n = 1, 2, \ldots)$ und λ ein Borelmaß auf $\mathfrak{B}$. Konvergiert dann die gleichgradig λ-stetige Folge (μ_n) vag gegen μ, so gilt* $\lim_{n \to \infty} \mu_n(B) = \mu(B)$ *für alle* $B \in \mathfrak{B}$.

BEWEIS. Da X separabel ist, kann nach einer Bemerkung in 2. λ o. B. d. A. als beschränkt vorausgesetzt werden. Sei also $B \in \mathfrak{B}$. Dann gibt es wegen der Regularität von λ zu jedem $\varepsilon' > 0$ eine offene Menge $U \subset X$ mit $U \supset B$ und eine kompakte Menge $K \subset X$ mit $K \subset B$, so daß $\lambda(U - K) < \varepsilon'$ ausfällt. Nach dem verallgemeinerten Satz von Urisohn existiert zu kompakter Menge K und offener Menge U mit $K \subset U$ eine Funktion $f \in \mathscr{C}^b$ mit

$0 \leq f \leq 1$ und $f(x) = 1$ für alle $x \in K$ sowie $f(x) = 0$ für alle $x \in U^c$. Wegen $\mu_n \xrightarrow{\text{vag}} \mu$ und Satz 3.2 existiert zudem zu vorgegebenem $\varepsilon > 0$ ein n_ε mit der Eigenschaft $|\mu_n(g) - \mu(g)| < \frac{\varepsilon}{3}$ für alle $g \in \mathscr{C}^b, n > n_\varepsilon$. Nach Voraussetzung der gleichgradigen λ-Stetigkeit von (μ_n) folgt aus $\lambda(U - K) < \varepsilon'$ sofort $\mu_n(U - K) < \frac{\varepsilon}{6}$ für alle $n = 1, 2, \ldots$. Also

$$|\mu_n(B) - \mu_n(K)| \leq \mu_n(U - K) < \frac{\varepsilon}{6}$$

für alle n. Mit Hilfe der Abschätzung

$$|\mu_n(B) - \mu(B)| \leq |\mu_n(f) - \mu_n(B)| + |\mu(f) - \mu(B)| + |\mu_n(f) - \mu(f)|$$

erkennt man, daß nur noch die ersten beiden Summanden auf der rechten Seite der Ungleichung abzuschätzen sind. Für den ersten Summanden ergibt sich

$$|\mu_n(f) - \mu_n(B)| \leq \left|\int\limits_{K \cup U^c} f d\mu_n - \mu_n(B)\right| +$$

$$+ \left|\int\limits_{K^c \cap U} f d\mu_n\right| \leq |\mu_n(K) - \mu_n(B)| + \mu_n(U - K) < \frac{\varepsilon}{3}.$$

Analoges folgt für den zweiten Summanden. Mithin erhalten wir für alle $n \geq n_\varepsilon$:

$$|\mu_n(B) - \mu(B)| < \varepsilon.$$

SATZ 5.2. *Seien $\mu, \mu_n \in \mathscr{M}^1_+ (n = 1, 2, \ldots)$ und λ ein Borelmaß auf $\mathfrak{B}$. Ist nun die Folge (μ_n) $(\varepsilon, \varepsilon')$-gleichgradig λ-stetig, so ist der vage Limes μ von (μ_n) $(\varepsilon, \varepsilon')$-$\lambda$-stetig.*

BEWEIS. Sei zunächst U eine offene Teilmenge von X mit $\lambda(U) < \varepsilon'$. Wegen der Regularität von μ existiert zu vorgegebenem $\varepsilon_1 > 0$ eine kompakte Menge $K_{\varepsilon_1} \subset U$ mit $\mu(U) < \mu(K_{\varepsilon_1}) + \varepsilon_1$. Nach dem verallgemeinerten Satz von Urisohn existiert ferner eine Funktion $f \in \mathscr{C}^b$ mit $0 \leq f \leq 1$, so daß $f(x) = 1$ auf K_{ε_1} und $f(x) = 0$ auf U^c gilt. $\mu_n \xrightarrow{\text{vag}} \mu$ impliziert nach Satz 3.2.

$$\mu(f) < \mu_n(f) + \varepsilon_2 < \mu_n(U) + \varepsilon_2 < \varepsilon + \varepsilon_2$$

zu vorgegebenem $\varepsilon_2 > 0$ für alle $n \geq n_{\varepsilon_2}$. Also haben wir

$$\mu(K_{\varepsilon_1}) < \mu(f) < \varepsilon + \varepsilon_2.$$

Da $\varepsilon_2 > 0$ beliebig vorgegeben war, folgt $\mu(K_{\varepsilon_1}) \leq \varepsilon$ und aus $\mu(U) < \mu(K_{\varepsilon_1}) + \varepsilon_1 \leq \varepsilon + \varepsilon_1$ schließlich $\mu(U) < \varepsilon$.

Sei nun $B \in \mathfrak{B}$ beliebig mit $\lambda(B) < \varepsilon'$ vorgegeben, $\varepsilon' > 0$ wie zuvor. Auf Grund der Regularität von λ existiert eine offene Menge $V \supset B$ mit $\lambda(V) < \lambda(B) + \varepsilon^*$ zu vorgegebenem $\varepsilon^* > 0$. Aus $\lambda(B) < \varepsilon'$ folgt $\lambda(V) < \varepsilon'$, womit die Voraussetzung des ersten Teils dieses Beweises erfüllt ist. Es ergibt sich $\mu(U) < \varepsilon$ und somit $\mu(B) \leq \mu(U) < \varepsilon$. ⌟

BEMERKUNG 5.1. Analoge Aussagen ergeben sich, wenn man anstelle der $(\varepsilon, \varepsilon')$-$\lambda$-Stetigkeit eine der beiden folgenden Voraussetzungen macht:

(a) μ_n ist λ-stetig für alle $n = 1, 2, \ldots$

(b) $\mu_n(B) < \alpha\lambda(B)$ für alle $B \in \mathfrak{B}$, gleichmäßig in $n = 1, 2, \ldots$ $(\alpha \in \mathbf{R}^1,$ $\alpha > 0)$.

Man sieht sofort, daß (a) aus (b) folgt.

Satz 5.1 und Satz 5.2 sind auf lokalkompakte (separable) Räume verallgemeinerte Resultate von Rivkind [13] und Kloss [10], die bisher nur für kompakte Räume X und ein spezielles Borelmaß λ bewiesen wurden. (Vgl. Bemerkung 8.2.1)

6. STRAFFE MENGEN VON W-VERTEILUNGEN

Aus der klassischen Theorie ist eine weitere zu (1) in Satz 4.1 äquivalente Aussage bekannt, die auf P. Lévy zurückgeht. Von Lévy stammt das heute in großer Allgemeinheit bekannte Resultat, daß die Menge der 1-dimensionalen Verteilungsfunktionen so metrisiert werden kann, daß der entstehende metrische Raum vollständig ist und die durch die Metrik erzeugte Topologie mit der Topologie der vagen Konvergenz übereinstimmt. Hiermit verknüpft sind Fragen über straffe Mengen von Verteilungsfunktionen, die wir jetzt für den Fall lokalkompakter Räume X im Hinblick auf gewisse Anwendungen in 8. beantworten wollen.

DEFINITION 6.1. Eine Teilmenge $\mathscr{M} \subset \mathscr{M}_+^b$ heißt *straff* auf X, wenn es zu jedem $\varepsilon > 0$ eine kompakte Teilmenge K_ε von X derart gibt, daß $\mu(K_\varepsilon^c) < \varepsilon$ gilt für alle $\mu \in \mathscr{M}$.

Der folgende wichtige Satz geht auf J. Dieudonné zurück:

SATZ 6.1. *Sei X lokalkompakt (mit abzählbarer Basis). Dann ist notwendig und hinreichend für die Straffheit einer Teilmenge $\mathscr{M} \subset \mathscr{M}_+^1$ die bezüglich der B-Topologie relative Kompaktheit von $\mathscr{M}$.*

BEWEIS nach [7], S. 220.

(A) Sei $\mathscr{M} \in \mathscr{M}_+^1$ straff, ferner $\{\mu_n\} \in \mathscr{M}$ mit $\mu_n \xrightarrow{B} \mu$ für $n \to \infty$ gegeben. Dann gilt für jede kompakte Teilmenge $K \subset X$: $\liminf \mu_n(K^c) \geq \mu(K^c)$. Also ist mit $\mathscr{M}$ auch $\overline{\mathscr{M}}$ (Abschluß im Sinne der B-Topologie) straff, und es folgt nach Kurzer Rechnung die Kompaktheit von $\overline{\mathscr{M}}$, also die B-Relativkompaktheit von $\mathscr{M}$.

(B) Sei nun $\mathscr{M} \in \mathscr{M}_+^1$ B-relativ kompakt. Es genügt, die Behauptung für B-kompaktes $\mathscr{M}$ zu beweisen. Sei (G_n) eine isotone Folge offener, relativ kompakter Mengen aus X, so daß $\bar{G}_n \subset G_{n+1}$ und $\bigcup_n G_n = X$ ist. Wäre $\mathscr{M}$ nicht straff, so gäbe es ein $\varepsilon > 0$, so daß für jedes $n = 1, 2, \ldots$ ein $\mu_n \in \mathscr{M}$ existiert mit $\mu_n(G_{n+1}^c) \geq \varepsilon$. Also existieren ein $l > n$ und eine Funktion $f_n \in \mathscr{C}^b$, so daß $\|f_n\| \leq 1$ und $f_n(x) = 0$ für $x \in G_{2n+1}^c \cap \bar{G}_{2l}$ mit $\mu_n(f_n) > \varepsilon/2$ gilt.

Wir definieren nun eine Folge $\{n_k\}$ natürlicher Zahlen durch $n_1 = 1$, $n_{k+1} = l_k + 1$. Zur Abkürzung werden im weiteren $\mu_k := \mu_{n_k}$ und $f_k := f_{n_k}$ gesetzt.

Sei S der Raum aller beschränkten Folgen reeller Zahlen, S^* der zugehörige Dualraum. Die Abbildung $T : S \to \mathcal{C}^b$ werde definiert durch

$$T(\xi_k) = \sum_k \xi_k f_k$$

für $(\xi_k) \subset (\mathbf{R}^1)$.

Für jedes $s \in S$ gilt $\| Ts \| \leq \| s \|$. Die Transponierte T^t von T definiert durch $< T^t \mu, s > = < \mu, Ts >$ bildet $\mathcal{M}_+^1$ in S^* ab. T^t ist B- und $\sigma(S^*, S)$-stetig (siehe [4], S. 101).* Also ist $T^t(\mathcal{M})$ $\sigma(S^*, S)$-kompakt in S^*. Sei B der Teilraum der Folgen mit nur endlich vielen Koordinaten $\neq 0$. B ist normseparabel, also kann man aus $(T^t \mu_k)$ eine Teilfolge $(T^t \mu_{k_j})_j$ auswählen, die punktweise auf B gegen $\mu_0 \in T^t(\mathcal{M})$ konvergiert.

Es folgt die Konvergenz der Folge $(T^t \mu_{k_j})_j$ gegen μ_0 im Sinne der $\sigma(S^*, S)$-Topologie. Nun ist die Menge $(T^t \mu_{k_j})_j$ straff in der Menge der natürlichen Zahlen. Sei s_k dasjenige Element aus S, dessen einzige nichtverschwindende Koordinate die k-te ist und den Wert 1 hat. Nach Definition ist $(T^t \mu_{k_j})\ (s_{k_j}) = \mu_{k_j}(f_{k_j}) \geq \varepsilon/2$. i. W. z. Straffheit der Menge $(T^t \mu_{k_j})_j$. ⌟

7. STOCHASTISCHE KONVERGENZ GRUPPENWERTIGER ZUFALLSVARIABLEN

Gemäß den Zusammenhängen in der klassischen Theorie zwischen stochastischer Konvergenz (auch Konvergenz nach W genannt), vager und Bernoullikonvergenz werde jetzt auf die Verallgemeinerung dieser Begriffsbildungen auf lokalkompakte Räume bzw. Gruppen eingegangen. Bekanntlich konvergiert in der klassischen Theorie eine Folge (X_n) reeller Zufallsvariablen genau dann stochastisch gegen X, wenn jede Teilfolge von (X_n) eine fast sicher gegen X konvergente Teilfolge besitzt. Aus der stochastischen Konvergenz folgt die Konvergenz der Folge der Verteilungsfunktionen der X_n auf den Stetigkeitsstellen der Grenzverteilungsfunktion. Schließlich hat man den Satz, daß die Umkehrung dieses Resultates nur dann gilt, wenn es sich um die Konvergenz der Verteilungsfunktionen gegen eine Konstante handelt.

Derartige Zusammenhänge gelten nun allgemeiner: Sei G eine lokalkompakte (separable) Gruppe. *Gruppenwertige Zufallsvariable* (oder kurz G-*Zufallsvariable*) auf einem W-Feld $(\Omega, \mathfrak{A}, P)$ heißen alle $\mathfrak{A}$-meßbaren Abbildungen von Ω in G.

Definition 7.1. Eine Folge (X_n) von G-Zufallsvariablen X_n *konvergiert stochastisch* gegen eine G-Zufallsvariable X, wenn für jede Borelsche Umgebung U_e der Einheit e gilt:

$$\lim_{n \to \infty} \mathbf{P}[X_n X^{-1} \notin U_e] = 0,$$

d.h. wenn für jede Umgebung U_e der Einheit und jedes $\varepsilon > 0$ die Beziehung $\mathbf{P}[X_n X^{-1} \in U_e] > 1 - \varepsilon$ für alle $n \geq n_0 = n_0(U_e, \varepsilon)$ gilt.

* Mit $\sigma(S^*, S)$-Topologie ist die durch S auf S^* definierte schwache Topologie gemeint. Analog hierzu wären die vage Topologie mit $\sigma(\mathcal{M}, \mathcal{K})$ und die B-Topologie mit $\sigma(\mathcal{M}^b, \mathcal{C}^b)$ zu bezeichnen.

Es ist klar, daß für eine gegen eine G-Zufallsvariable stochastisch konvergierende Folge von G-Zufallsvariablen (X_n) auch die Cauchyfolge $(X_n X_m^{-1})$ stochastisch gegen e konvergiert.* Die Umkehrung dieses Satzes wird in [8] auf S. 108 bewiesen.

SATZ 7.1. *Konvergiert die Cauchyfolge* $(X_n X_m^{-1})_{n,m}$ *stochastisch gegen* e *für* $n, m \to \infty$, *so auch die Folge* (X_n).

BEWEIS. Wir wählen eine Folge (ε_k) positiver reeller Zahlen mit $\sum\limits_k \varepsilon_k < \infty$ sowie eine antitone Folge (U_k) von Umgebungen der Einheit mit $U_k U_k \subset \subset U_{k-1}$ und $\bigcap\limits_k U_k = \{e\}$. Es sei n_k so gewählt, daß

$$\mathbf{P}[X_{n_k} X_{n_{k+1}}^{-1} \in U_k] > 1 - \varepsilon_k$$

erfüllt ist. Definiert man die Mengen $S_l := \bigcap\limits_{k=l}^{\infty} [X_{n_k} X_{n_{k+1}}^{-1} \in U_k]$ und $S = \bigcup\limits_{l=1}^{\infty} S_l$, so erhält man sofort

$$\mathbf{P}(S_l^c) \leq \sum_{k=l}^{\infty} \mathbf{P}[X_{n_k} X_{n_{k+1}}^{-1} \notin U_k] \leq \sum_{k=l}^{\infty} \varepsilon_k$$

und weiterhin

$$\mathbf{P}(S_l) \geq 1 - \sum_{k=l}^{\infty} \varepsilon_k ,$$

woraus für $l \to \infty$ $\mathbf{P}(S) = 1$ folgt. Also ist $(X_{n_k})_k$ eine Cauchyfolge; denn zu vorgegebener Umgebung U der Einheit und beliebigem $\varepsilon > 0$ kann man eine natürliche Zahl p so finden, daß $U_p \subset U$ und für $p < k < l$ $X_{n_k} X_{n_k}^{-1} \in \in U$ gilt. Letzteres folgt mit Hilfe von $U_k^2 \subset U_{k-1}$ aus der Inklusion

$$X_{n_k} X_{n_l}^{-1} = X_{n_k} X_{n_{k+1}}^{-1} X_{n_{k+1}} \cdot \ldots \cdot X_{n_{l-1}} X_{n_l}^{-1} \in$$

$$\in U_k U_{k+1} \cdot \ldots \cdot U_{l-1} \subset U_{k-1} \subset U_p \subset U.$$

Aufgrund der Vollständigkeit lokalkompakter Gruppen (vgl. [6] S. 40), existiert ein X, so daß die Folge $(X_{n_k})_k$ gegen X stochastisch konvergiert. Da zudem $(X_n X_{n_k}^{-1})$ gegen e stochastisch konvergiert, erhalten wir wegen $X_n X^{-1} = X_n X_{n_k}^{-1} X_{n_k} X^{-1}$ die Behauptung. ⌟

SATZ 7.2. *Seien* $X, X_1, X_2, \ldots$ G-*Zufallsvariable mit den* W-*Verteilungen* $\mu, \mu_1, \mu_2, \ldots$. *Dann folgt aus der stochastischen Konvergenz der Folge* (X_n) *gegen* X *die vage Konvergenz der Folge ihrer* W-*Verteilungen* (μ_n) *gegen* μ.

BEWEIS. Nach Satz 4.1 genügt es zu zeigen, daß aus der stochastischen Konvergenz der X_n gegen X die Konvergenz der μ_n gegen μ auf den μ-quadrierbaren Borelschen Mengen $B \subset G$ folgt. Sei $B \in \mathfrak{B}$ μ-quadrierbar. Dann spalten wir auf:

$$\mu_n(B) - \mu(B) = \mathbf{P}[X_n \in B] - \mathbf{P}[X \in B] = \mathbf{P}[(X_n X^{-1}) X \in B] - \mathbf{P}[X \in B] =$$

$$\mathbf{P}[(X_n X^{-1}) X \in B] \cap [X \in B^\varepsilon] + \mathbf{P}[(X_n X^{-1}) X \in B] \cap [X \notin B^\varepsilon] - \mathbf{P}[X \in B]$$

* e ist hier als konstante G-Zufallsvariable verstanden.

für alle μ-quadrierbaren $B \in \mathfrak{B}$ und jedes $B^\varepsilon \in \mathfrak{B}$. Wegen der Regularität von μ existieren zu vorgegebenem $\varepsilon > 0$ eine offene Menge $U \subset G$ mit $U \supset \bar{B}$ und eine kompakte Menge $K \subset G$ mit $K \subset \mathring{B}$, so daß $\mu(U - K) < \varepsilon$ wird. Nun kann man eine symmetrische Umgebung U_e der Einheit so wählen, daß $U_e K \subset B$ gilt. Man wähle ferner n_ε so, daß nach Voraussetzung $\mathbf{P}[X_n X^{-1} \notin U_e] < \varepsilon$ ist für alle $n \geq n_e$. Nach Definition $B^\varepsilon := U_e K$ folgt einmal

$$\mathbf{P}\,[(X_n X^{-1}) X \in B] \cap [X \notin B^\varepsilon] < 2\varepsilon\,.$$

Wegen $\mathbf{P}[(X_n X^{-1}) X \in B] \cap [X \in B^\varepsilon] \geq \mathbf{P}\,[(X_n X^{-1}) X \in B] \cap [X \in K] \geq$ $\geq \mathbf{P}\,[X \in K] - \varepsilon$ folgt schließlich für alle $n \geq n_\varepsilon$ die Behauptung

$$|\,\mu_n(B) - \mu(B)\,| < 2\varepsilon. \quad \lrcorner$$

Bemerkung 7.1. Die Umkehrung dieses Satzes gilt nicht, wie folgendes Beispiel zeigt. Man nehme $G = \mathbf{R}^1, \Omega = [0, 1], \mathfrak{A} = \mathfrak{B}^1 \cap \Omega, P =$ Lebesguemaß. Definiere

$$A_n = \left[0, \frac{1}{2^n}\right] \cup \left[\frac{2}{2^n}, \frac{3}{2^n}\right] \cup \cdots \cup \left[\frac{2^n - 2}{2^n}, \frac{2^n - 1}{2^n}\right]$$

und $X_n = \chi_{A_n}$ für alle $n = 1, 2, \ldots$. Dann ergibt sich $\mu_{X_n} = \frac{1}{2}\varepsilon_0 + \frac{1}{2}\varepsilon_1$ für alle $n = 1, 2, \ldots$. Obgleich $\mu_n(f) \to \mu(f)$ für alle $f \in \mathscr{K}$ gilt, erhält man wegen

$$\mathbf{P}\left[|\,X_n - X\,| \geq \frac{1}{2}\right] = \frac{1}{2}$$

keine stochastische Konvergenz der Folge (X_n).

Hingegen kann man folgenden Satz beweisen, der im Speziallfall die Äquivalenz von stochastischer und vager Konvergenz beinhaltet.

Satz 7.3. *Sei (X_n) eine Folge von G-Zufallsvariablen, (μ_n) die Folge ihrer W-Verteilungen. Dann sind stochastische Konvergenz der X_n gegen e und vage Konvergenz der μ_n gegen ε_e äquivalent.*

Beweis. Wegen Satz 7.2 genügt es zu zeigen, daß aus der vagen Konvergenz gegen das Punktmaß im Einselement die stochastische Konvergenz der Folge (X_n) gegen e folgt. Hierzu betrachten wir eine beliebige Borelsche Umgebung U_e der Einheit e. Es seien $g, f \in \mathscr{K}$ mit $g \leq \chi_{U_e} \leq f$ und $g(e) = f(e) = 1$ gegeben. Dann folgt sofort $g \circ X_n \leq \chi_{U_e} \circ X_n \leq f \circ X_n$. Definieren wir $A_n := [X_n \notin U_e]$, so schreiben wir weiter $\chi_{U_e} \circ X_n = \chi_{A_n}$, und es ergibt sich

$$\mu_n(g) \leq 1 - \mathbf{P}(A_n) \leq \mu_n(f)\,.$$

Nach Voraussetzung gilt:

$$\lim_{n \to \infty} \mu_n(g) = \varepsilon_e(g) = 1$$

für alle $g \in \mathscr{K}$ und analog für alle $f \in \mathscr{K}$. Also erhält man $\lim_{n \to \infty} \mathbf{P}\,(A_n) = 1$. $\lrcorner$

8. ZWEI GRENZWERTSÄTZE FÜR W-VERTEILUNGEN AUF LOKALKOMPAKTEN GRUPPEN

8.1. Im folgenden sei nun wieder eine lokalkompakte (separable) Gruppe G zugrunde gelegt. Es ist bekannt, daß $\mathcal{M}^b$ Banachalgebra ist, wenn man die Faltung zweier Maße $\mu, \nu \in \mathcal{M}^b$ als Multiplikation einführt. In formaler Verallgemeinerung der klassischen Theorie definiert man

$$\mu * \nu(B) = : \mu\nu(B) = : \int \mu(Bx^{-1})\, \nu(dx) = \int \nu(x^{-1}B)\, \mu(dx)$$

für alle $B \in \mathfrak{B}$.

Die Abbildung $(\mu, \nu) \to \mu\nu$ von $\mathcal{M}^1_+ \times \mathcal{M}^1_+$ in $\mathcal{M}^1_+$ ist simultan Bernoulli-stetig in beiden Komponenten. Der Vollständigkeit wegen führen wir den Beweis des folgenden Satzes nach [8], S. 188 durch.

SATZ 8. 1. 1. *Seien* (μ_n) *und* (ν_n) *Folgen aus* $\mathcal{M}^1_+$ *mit den* B-*Limiten* μ *und* ν. *Dann gilt* $\mu_n\nu_n \xrightarrow{B} \mu\nu$.

BEWEIS. Nach Definition der Faltung gilt für alle $f \in \mathcal{C}^b$ mit $t_n(y) = \int f(xy)\mu_n(dx)$ die Gleichung

$$\int f(z)\, \mu_n\nu_n(dz) = \int\int f(xy)\, \mu_n(dx)\, \nu_n(dy) = \int t_n(y)\, \nu_n(dy)\,.$$

Definiert man entsprechend $t(y) = \int f(xy)\mu(dx)$ für $f \in \mathcal{C}^b$, so gilt auf Grund der Voraussetzung für alle $y \in G$:

$$t_n(y) \to t(y).$$

Da jede B-konvergente Folge in $\mathcal{M}^1_+$ relativ kompakt ist, ist (μ_n) straff nach Satz 6.1, und es gibt also zu jedem $\varepsilon > 0$ eine kompakte Menge $K_\varepsilon \subset G$ derart, daß $\mu_n(K_\varepsilon) \geq \geq 1 - \varepsilon$ ist. Es folgt sofort

$$|\, t_n(y) - t_n(y_0)\,| \leq 2\varepsilon \sup_{x \in G} |\, f(x)\,| + \max_{x \in K} |\, f(xy) - f(xy_0)\,|\,.$$

Also ist (t_n) gleichgradig stetig auf jeder kompakten Teilmenge von G, und daher konvergiert die Folge (t_n) gleichmäßig in K gegen t. Mithin ergibt sich die Behauptung so:

$$\int f(z)\, \mu_n\nu_n(dz) - \int f(z)\, \mu\nu(dz) = \int_K \big(t_n(y) - t(y)\big)\, \nu_n(dy) +$$

$$+ \int_{K^c} \big(t_n(y) - t(y)\big)\, \nu_n(dy) + \int t(y)[\nu_n(dy) - \nu(dy)] \to 0$$

für $n \to \infty$.

8.2. Auf jeder lokalkompakten Gruppe G existiert ein bis auf eine multiplikative Konstante eindeutiges reguläres Borelmaß ω mit den folgenden Eigenschaften:

(1) $\omega \not\equiv 0$ auf $\mathfrak{B}$ oder äquivalenterweise
(1′) $\omega(U) > 0$ für alle offenen nichtleeren Mengen $U \subset G$
(2) ω ist *linksinvariant*, d.h. es gilt $\omega(xB) = \omega(B)$ für alle $B \in \mathfrak{B}$, $x \in G$.

In gleicher Weise definiert man das rechtsinvariante Maß. Die so definierten Borelmaße heißen *Haarsche Maße* auf G.

Das Haarsche Maß ist normierbar zu einem W-Maß durch die Festsetzung $\omega(G) = 1$, falls $\omega(G)$ endlich ist, und dies ist genau dann der Fall, wenn G kompakt ist. Zu jeder kompakten Untergruppe H existiert genau ein Maß $\omega_H \in \mathscr{M}^1_+$ mit den Eigenschaften $\mathrm{Tr}\,\omega_H = H$ und $\varepsilon_x \omega_H = \omega_H \varepsilon_x$ für alle $x \in H$.

Im Anschluß an Satz 5.2 machen wir noch folgende

BEMERKUNG 8.2.1. Es folgt aus der $(\varepsilon, \varepsilon')$-$\omega$-Steigkeit von $\mu \in \mathscr{M}^1_+$ die $(\varepsilon, \varepsilon')$-gleichgradige ω-Stetigkeit von (μ^n).

Dies sieht man für den Fall $n = 2$ so : Sei $B \in \mathfrak{B}$. Wegen $\omega(B) = \omega(x^{-1}B)$ folgt aus $\omega(B) < \varepsilon'$ sofort $\mu(x^{-1}B) < \varepsilon$ nach Voraussetzung und damit

$$\mu^2(B) = \int \mu(x^{-1}B)\, \mu(dx) < \varepsilon$$

für alle $B \in \mathfrak{B}$, $x \in G$. Der Rest folgt bei Anwendung vollständiger Induktion.

In gleicher Weise ergibt sich aus der ω-Stetigkeit von $\mu \in \mathscr{M}^1_+$ die gleichgradige ω-Stetigkeit von (μ^n).

8.3. Die Theorie der im Sinne der Faltung unzerlegbaren W-Verteilungen hat K. R. Parthasarathy, R. R. Rao und S. R. S. Varadhan [11] zu Verallgemeinerungen geführt, die Anlaß zu den folgenden Betrachtungen geben.

DEFINITION 8.3.1. Sei $\mu \in \mathscr{M}^1_+$, $x \in G$. Dann heißt jedes W-Maß der Gestalt $\mu\, \varepsilon_x$ *x-Rechtstranslation* von μ. Analog definiert man die *x-Linkstranslation*. Ein Maß $\mu \in \mathscr{M}^1_+$ heißt *Translation* eines Maßes $\nu \in \mathscr{M}^1_+$, wenn μ entweder x-Rechts- oder Linkstranslation von ν ist.

BEMERKUNG 8.3.1. Die Bezeichnung x-Rechtstranslation von μ für $\mu\varepsilon_x$ rechtfertigt sich aus der Berechnung

$$\mu\varepsilon_x(B) = \mu(Bx^{-1})$$

für alle $B \in \mathfrak{B}$, $x \in G$.

DEFINITION 8.3.2. Eine Menge $\mathscr{N} \subset \mathscr{M}^1_+$ heißt *translationskompakt*, wenn es zu jeder Folge $(\mu_n) \in \mathscr{N}$ eine Folge $(\nu_n) \subset \mathscr{M}^1_+$ gibt mit den folgenden Eigenschaften:

(a) ν_n ist Translation von μ_n für $n = 1, 2 \ldots$

(b) (ν_n) ist relativ kompakt.

SATZ 8.3.1. *Seien (λ_n), (μ_n) und (ν_n) Folgen in $\mathscr{M}^1_+$ mit $\lambda_n = \mu_n \nu_n$ für alle $n = 1, 2, \ldots$. Dann folgt aus der Relativkompaktheit der Folge (λ_n) im Sinne der Bernoulli-Topologie die Translationskompaktheit von (μ_n) und (ν_n).*

Umgekehrt impliziert die Relativkompaktheit von (μ_n) und (ν_n) die Relativkompaktheit von (λ_n).

BEWEIS. (A) Wir führen den Nachweis der ersten Aussage gemäß [11] in drei Schritten:

1. Gezeigt wird: Sind (λ_n) und (μ_n) relativ kompakt, so ist es auch die Folge (ν_n). Nach Voraussetzung und Satz 6.1 gibt es zu jedem $\varepsilon > 0$ kompakte Mengen K_ε, K'_ε aus G, so daß $\lambda_n(K_\varepsilon) > 1 - \varepsilon$ und $\mu_n(K'_\varepsilon) > 1 - \varepsilon$ für alle $n = 1, 2, \ldots$ gilt. Es ist also

$$1 - \varepsilon < \lambda_n(K_\varepsilon) = \int \nu_n(x^{-1}K_\varepsilon)\, \mu_n(dx) \leq \int\limits_{K'_\varepsilon} \nu_n(x^{-1}K_\varepsilon)\, \mu_n(dx) + \varepsilon\,,$$

also

$$\int\limits_{K'_\varepsilon} \nu_n(x^{-1}K_\varepsilon)\, \mu_n(dx) > 1 - 2\varepsilon\,.$$

O.B.d.A. können wir $K'_\varepsilon = K_\varepsilon$ setzen; denn andernfalls nehme man $K_\varepsilon \cup K'_\varepsilon$. Also existiert ein $x_n \in K_\varepsilon$ so, daß $\nu_n(x_n^{-1}K_\varepsilon) > 1 - 2\varepsilon$ und weiterhin $\nu_n(K_\varepsilon^{-1}K_\varepsilon) > 1 - 2\varepsilon$ ausfällt für alle $n = 1, 2, \ldots$. Aus der Kompaktheit von $K_\varepsilon^{-1}K_\varepsilon$ sowie aus der Unabhängigkeit von n folgt Teilbehauptung 1.

2. Gezeigt wird: Unter den Voraussetzungen des Satzes ist $(\mu_n\, \varepsilon_{x_n})$ relativ kompakt. Man wähle eine Folge (ε_k) positiver reeller Zahlen mit $\sum\limits_k \varepsilon_k < \infty$. Da (λ_n) relativ kompakt ist, gibt es nach Satz 6.1 zu jedem ε_k eine kompakte Menge $K_k \subset G$, so daß $\lambda_n(K_k) > 1 - \varepsilon_k$ gilt für alle $k, n = 1, 2, \ldots$. Weiterhin nehme man eine Folge (ε'_k) positiver reeller Zahlen mit $\varepsilon'_k \downarrow 0$, so daß $\sum\limits_k \frac{\varepsilon_k}{\varepsilon'_k} \leq \frac{1}{2}$ erfüllt ist. Sei

$$A_{n_k} = \{x : \mu_n(K_k x^{-1}) > 1 - \varepsilon'_k\}$$

und $B_n = \bigcap\limits_k A_{n_k}$. Wir zeigen: $B_n \neq \varnothing$. Dies folgt aus:

$$1 - \varepsilon_k < \lambda_n(K_k) = \int\limits_{A_{n_k}} \mu_n(K_k x^{-1})\, \nu_n(dx) + \int\limits_{A^c_{n_k}} \mu_n(K_k x^{-1})\, \nu_n(dx) \leq$$

$$\leq \nu_n(A_{n_k}) + (1 - \varepsilon'_k)\, \nu_n(A^c_{n_k})$$

und

$$\nu_n(A^c_{n_k}) \leq \frac{\varepsilon_k}{\varepsilon'_k}\,,$$

also

$$\nu_n(B^c_n) = \nu_n\Big(\bigcup_k A^c_{n_k}\Big) \leq \sum \frac{\varepsilon_k}{\varepsilon'_k} \leq \frac{1}{2}\,.$$

Sei nun $x_n \in B_n$. Dann haben wir

$$\mu_n(K_k x^{-1}) > 1 - \varepsilon'_k$$

für alle $k, n = 1, 2, \ldots$, woraus Teilbehauptung 2. folgt.

3. Der Rest des Beweises ergibt sich so: Wir definieren $\alpha_n := \mu\varepsilon_{x_n}$ und $\beta_n := \varepsilon_{x_n^{-1}}\, \nu_n$. Es ist $\lambda_n = \alpha_n\, \beta_n$. (λ_n) und nach 2. dieses Beweises auch (α_n) sind relativ kompakt, also ist nach 1. dieses Beweises (β_n) relativ kompakt. Da α_n Translation von μ_n und β_n Translation von ν_n für alle $n = 1, 2, \ldots$

ist, ergibt sich die Forderung (a) von Definition 8.3.2 für (μ_n) und (ν_n), womit der Satz bewiesen ist.

(B) Für beliebige kompakte Mengen K_1, $K_2 \subset G$ folgt

$$\mu_n\nu_n(K_1K_2) \geq \mu_n(K_1)\,\nu_n(K_2)$$

für alle $n = 1, 2, \ldots$ aus der Ungleichung

$$\chi_{K_1K_2}(xy) \geq \chi_{K_1}(x)\,\chi_{K_2}(y)$$

durch Integration bezüglich μ_n und ν_n. Zweimalige Anwendung von Satz 6.1 ergibt die Behauptung. ⌟

BEMERKUNG 8.3.2. Im Sinne der vagen Topologie erhält man die Relativkompaktheit der Folge (λ_n) bereits für je zwei Folgen (μ_n) und (ν_n) aus $\mathscr{M}_+^1$ mittels der Abschätzung

$$\sup_n |\lambda_n(f)| \leq \sup_n \|\lambda_n\|\,\|f\| = \|f\| < \infty$$

für alle $f \in \mathscr{K}$ und dem Satz, daß jede beschränkte Teilmenge von $\mathscr{M}$ relativ kompakt ist.

LITERATUR

1. BAUER, H., Zur Theorie des Riemannintegrals in lokalkompakten Räumen· Bayer. Akad. der Wiss. München 187–208 (1955).
2. BAUER, H., Über die Beziehungen einer abstrakten Theorie des Riemannintegrals zur Theorie Radonscher Maße. *Math. Zeitschr.* **6,** 448–482 (1956).
3. BAUER, H., Wahrscheinlichkeitstheorie und Mathematische Statistik. Vorlesung, gehalten an der Universität Hamburg (1962/63).
4. BOURBAKI, N., Espaces vectoriels topologiques. Chapitres III–V, Hermann et Cie (1955).
5. BOURBAKI, N., Intégration, Chapitres I–IV. Hermann et Cie (1960).
6. BOURBAKI, N., Topologie générale, Chapitres III–IV. Hermann (1960).
7. LECAM, L., Convergence in Distribution of Stochastic Processes. University of California Press. Berkeley and Los Angeles (1957).
8. GRENANDER, U., Probabilities on Algebraic Structures. Almquist and Wiksell, John Wiley and Sons (1963).
9. LOOMIS, L., Linear Functionals and Content. *Am. Journ. of Math.* **76,** 168–182 (1954).
10. KLOSS, M. B., Limiting Distributions on Bicompact Abelian Groups, übersetzt in *Theory of Prob. and its Appl.* **VI,** 361–389 (1961).
11. PARTHASARATHY, K. R., RAO, R. R., VARADHAN, S. R. S., On the Category of Indecomposable Distributions on Topological Groups. *TAMS* **102,2,** 200–217 (1962).
12. RAO, R. R., Some Problems in Probability Theory. Thesis submitted to Calcutta University (1960).
13. RIVKIND, YA. I., Limit Theorem of Probability Theory on Compact Topological Groups, übersetzt in *Selected Translations in Math. Stat. and Prob.* **2,** 175–181 (1962).

STATISTISCHE STICHPROBEN VON LANGGESTRECKTEN MATERIALEN

von

J. Křepela und P. Mandl

Institut für Informationstheorie und Automatisierung
der Tschechoslowakischen Akademie der Wissenschaften
Prag, Tschechoslowakei

Der vorliegende Beitrag behandelt die Kontrolle von Materialen, deren Qualität davon abhängt, ob eine gewisse Größengrenze überschritten worden ist oder nicht. Trotzdem es sich um eine allgemeine Methode handelt, werden wir uns bei ihrer Darlegung mit folgendem konkreten Beispiel behelfen. Auf einen Leitungsdraht wird eine Isolationsschicht aufgetragen. Die Dicke dieser Schicht ist im Verlaufe der Produktion zufälligen Änderungen unterworfen. Wir setzen voraus, daß nur die Größe der Schwankung der Dicke in der Längerichtung von praktischer Bedeutung ist. Es liegt also ein Zufallsprozeß $\{X(t)\}$ vor, wobei $X(t)$ die Dicke der Isolationsschicht in der Entfernung t vom Anfang des Drahtes bedeutet. Die Qualität der Isolierung wird natürlich nicht eindeutig durch die Dicke der Schicht bestimmt. Es können z. B. Blasen vorkommen, welche die Isolationsfähigkeit vermindern. Für den Anfang werden wir die Existenz solcher Blasen ausschließen. Dieser Fall läßt sich auf eine natürliche Weise als ein zusätzlicher Poissonscher Prozeß im vorgeschlagenen Kontrollverfahren in Betracht ziehen.

In den meisten praktischen Fällen wäre ein, auf der genauen Festlegung der Stellen von ungenügender Isolationsschicht beruhendes Verfahren zu kostspielig. Man nimmt also bei der Kontrolle des isolierten Drahtes einen Abschnitt desselben und prüft, ob bei einer gegebenen Spannung die Isolation durchschlagen wird oder nicht. Nach der Voraussetzung erfolgt der Durchschlag genau dann, wenn $\{X(t)\}$ in dem Abschnitt unter eine Grenze Z sinkt. Das ist eine Situation, welche man nicht nur bei Leitungsdraht sondern auch bei anderen langgestreckten Materialen auffindet.

Im Punkt 1 werden einige Resultate über die Erst-Passage Zeiten der Zufallsprozesse angegeben. Besondere Aufmerksamkeit wird dem Uhlenbeckschen Prozeß, welcher vom Professor J. Hájek als geeignetes Modell für die Schwankungen der Qualität in langgestreckten Materialen vorgeschlagen wurde, gewidmet. Im Punkt 2 wird das Problem der Kontrolle der Isolationsschicht mit Hilfe des Uhlenbeckschen Modells formuliert und unter der Voraussetzung der Richtigkeit der im Punkte 1 angegebenen Annäherungen auf ein Problem überführt, das in gewissem Sinne von den Eigenschaften des unterliegenden mathematischen Modells unabhängig ist. Eine Methode des Zusammensetzens von Stichprobenplänen wird im Punkt 3 dargestellt und im Punkt 4 an einem Beispiel beleuchtet.

1. Es sei $\{Y(t)\}$, $t \geq 0$, ein stetiger eindimensionaler Zufallsprozeß. Es sei Z eine Zahl, $\tau(Z)$ die Zeit, in welcher $\{Y(t)\}$ zum erstenmal einen Wert kleiner als Z angenommen hat. Wir bezeichnen $\mathbf{E}\,\tau(Z) = m(Z)$. Für eine breite Klasse von stationären Zufallsprozessen, die sicher fast alle praktisch interessanten stationären Prozesse umfaßt, gilt

$$\lim_{Z \to -\infty} \mathbf{P}\big(\tau(Z) \geq m(Z)t\big) = \mathrm{e}^{-t}.* \tag{1}$$

Setzen wir voraus, daß $\{Y(t)\}$ ein homogener Diffusionsprozeß mit Übergangswahrscheinlichkeitsdichte $f^{(t)}(x, y)$ ist. $f^{(t)}(x, y)$ erfülle die Fokker–Plancksche Gleichung

$$\frac{\partial}{\partial t} f^{(t)}(x, y) = \frac{\partial^2}{\partial y^2}\left[a(y)\, f^{(t)}(x, y)\right] - \frac{\partial}{\partial y}\left[b(y)\, f^{(t)}(x, y)\right].$$

Die Funktionen $a(y)$, $b(y)$ seien stetig, $a(y) > 0$. Wir setzen

$$W(x) = \exp \int_0^x b(s)\, a(s)^{-1}\, ds\,.$$

Unter der Voraussetzung

$$M = \int_{-\infty}^{\infty} a(x)^{-1}\, W(x)\, dx < \infty$$

hat man

$$\lim_{Z \to -\infty} \mathbf{P}\big(\tau(Z) \geq m(Z, x)\, t \,|\, x(0) = x\big) = \mathrm{e}^{-t}$$

mit

$$m(Z, x) = \mathbf{E}\{\tau(Z) | x(0) = x\} = \int_Z^x W(s)^{-1} \int_s^{\infty} a(y)^{-1}\, W(y)\, dy\, ds^{**}\,. \tag{2}$$

Nimmt man für die Dichtefunktion von $Y(0)$ die Funktion

$$g(y) = M^{-1}\, a(y)^{-1}\, W(y), \tag{3}$$

so wird $\{Y(t)\}$ stationär und man hat

$$m(Z) = M^{-1} \int_Z^{\infty} W(s)^{-1} \Big(\int_s^{\infty} a(y)^{-1}\, W(y)\, dy\Big)^2 ds\,.$$

Ist $m(Z)$ nicht endlich, so muß man in (1) $m(Z)$ durch (2) ersetzen.

Ein Diffusionsprozeß ist ein Uhlenbeckscher Prozeß, wenn $a(y) \equiv \beta\sigma^2$, $b(y) \equiv -\beta(y - \mu)$ ist, wo $\sigma > 0$, $\beta > 0$. Die stationäre Verteilung (3) ist eine Normalverteilung mit Mittelwert μ und Streuung σ^2. Der stationäre Prozeß ist dann ein Gaußscher Prozeß mit der Korrelationsfunktion $\exp -\beta|t|$, und es gilt

$$m(Z) = \beta^{-1} \varphi\big((Z - \mu)/\sigma\big)\,. \tag{4}$$

* Siehe Volkonskij, V. A., Rozanov, Ju. A., *Teorija verojatnostej i prim.* **IV**, 186–207 (1959).

** Siehe Mandl, P., *Transactions of the Third Prague Conf. on Inform. Th. etc.* S. 475–498.

Dabei ist

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{y^2/2} \left(\int_y^\infty e^{-s^2/2}\, ds \right)^2 dy .$$

Die Werte von $\varphi(x)$ in Tabelle I sind einer ausführlicheren Tafel entnommen.*

TABELLE I

$-x$	$\varphi(x)$	$\psi(-x)$	$-x$	$\varphi(x)$	$\psi(-x)$
2,0	9,69	3,75	4,0	2 017,03	1 006,81
2,5	26,08	11,76	4,5	14 730,82	7 363,59
3,0	85,84	41,48	5,0	140 739,43	70 367,80
3,5	363,12	179,98			

Die mathematische Erwartung der Zeit, in welcher der stationäre Uhlenbecksche Prozeß zum erstenmal einen Wert annimmt, der mindestens um Z, $Z > 0$ vom Mittelwert μ abweicht, ist $\beta^{-1}\psi(Z/\sigma)$ mit

$$\psi(x) = \sqrt{\frac{2}{\pi}} \int_0^x e^{y^2/2} \left(\int_0^y e^{-s^2/2}\, ds \right)^2 dy .$$

Tabelle I enthält einige Werte von $\psi(x)$.

2. Kehren wir jetzt zum Zufallsprozeß $\{X(t)\}$, welcher die Dichte der Isolationsschicht in der Partie vom Leitungsdraht beschreibt, zurück. Wir setzen voraus, daß $\{X(t)\}$ ein stationärer Uhlenbeckscher Prozeß ist, dessen Parameter σ, β von Posten zu Posten verhältnismäßig stabil sind, der Parameter μ aber kann variieren und ist also für die Qualität bestimmend.

Die Entscheidung über die Zulässigkeit der Posten soll nach einer Stichprobe von Abschnitten des Drahtes genommen werden. Es sei $p(\mu, x)$ die Wahrscheinlichkeit, daß in einem gegebenen Abschnitt die Isolationsschicht durchschlagen wird, d. h. ihre Dicke unter die Grenze Z sinkt. Man hat nach (1), (4)

$$p(\mu, x) \sim 1 - \exp\{ -\beta\varphi\big((Z-\mu)/\sigma\big)^{-1} x\} . \tag{5}$$

(Zur Übereinstimmung der Annäherung in (5) siehe die gerade zitierte Arbeit). Ein einfacher Stichprobenplan wird also aus der Stichprobengröße n, aus dem Kriterium c und aus der Abschnittslänge x bestehen.

Wir setzen

$$k = \beta\varphi\big((Z-\mu)/\sigma\big)^{-1} . \tag{6}$$

* Siehe MANDL, P., *Aplikace matematiky* 141–148 **7**, (1962).

k ist eine abnehmende Funktion von μ. In die Formel (6) wird man in der Praxis für β, σ ihre Schätzwerte, die sich aus der Beobachtung der Korrelation und der Streuung der Dicke der Isolationsschicht ergeben, einsetzen. Nimmt man k als den Parameter für die Qualität der Posten, so ist man zu der Aufgabe gekommen einen Stichprobenplan zusammenzusetzen, wenn die Ausschußwahrscheinlichkeit von einem Parameter x abhängt. Die Abhängigkeit ist

$$p = 1 - \mathrm{e}^{-kx} \tag{7}$$

und die Wahl der Größe von x liegt uns ob.

Aus der Allgemeinheit der Beziehung (1) folgt, daß die Formulierung der Aufgabe nicht von der Annahme, daß $\{X(t)\}$ ein Uhlenbeckscher Prozeß ist, abhängt. Auch die Stabilität von σ, β ist nicht entscheidend. Wenn in der Isolation Blasen in der Form eines, von $\{X(t)\}$ unabhängigen Poissonschen Prozesses mit der Intensität λ vorkommen, so ist die Wahrscheinlichkeit einer Blase im Abschnitt gleich $1 - e^{-\lambda x}$ und man ist zu der Formel

$$p = 1 - \mathrm{e}^{-(k+\lambda)x} \tag{7'}$$

gelangt.

3. Beim Zusammenstellen eines Stichprobenplanes für statistische Annahme von Posten des Leitungsdrahtes gehen wir von der häufig benutzten Voraussetzung aus, daß Posten mit $k \leqq k_0$ mit einer Wahrscheinlichkeit größer als $1 - \alpha_0$, Partien mit $k \geqq k_1$ mit einer Wahrscheinlichkeit kleiner als α_1 angenommen werden sollen.

Es sei $L(n, c, p)$ die Annahmekennlinie des Stichprobenplanes (n, c). Der Stichprobenplan ist anwendbar, wenn sich eine Abschnittslänge x derart finden läßt, daß man

$$\begin{aligned} L(n, c, 1 - \mathrm{e}^{-k_0 x}) &\geqq 1 - \alpha_0 \\ L(n, c, 1 - \mathrm{e}^{-k_1 x}) &\leqq \alpha_1 \end{aligned} \tag{8}$$

hat.

Ist

$$L(n, c, p_0) = 1 - \alpha_0\,, \qquad L(n, c, p_1) = \alpha_1\,,$$

so muß

$$\mathrm{e}^{-k_0 x} \geqq 1 - p_0\,, \qquad \mathrm{e}^{-k_1 x} \leqq 1 - p_1\,, \tag{9}$$

daher

$$\frac{k_1}{k_0} \geqq k(n, c) = \frac{\ln(1 - p_1)}{\ln(1 - p_0)} \tag{10}$$

gelten. Ist (10) erfüllt, so kann man (9), d. h. (8) mit

$$x = -\, k_1^{-1} \ln(1 - p_1) \tag{11}$$

erzielen.

Die Ausgaben für das Kontrollverfahren seien durch eine Funktion $C(n, y)$ gegeben. $y = nx$ ist die Menge des bei der Kontrolle verbrauchten Materials. Man wählt einen Stichprobenplan, für welchen $C(n, y)$ minimal wird. Dabei kann die folgende Tabelle II nützlich sein. Sie enthält für

$\alpha_0 = \alpha_1 = 0{,}05$, $C = 0, \ldots, 10$, $n - C = 1, \ldots, 10$ die Werte von $k(n, c)$ und von

$$\eta(n, c) = -n \ln(1 - p_1) .$$

In der letzten Spalten befinden sich die Grenzwerte für $n = \infty$. Die Anwendung der Tabelle wird an einem Beispiel erklärt.

TABELLE II*

c \ $n-c$	1	2	3	4	5	6	7	8	9	10	∞
0	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00	58,4 3,00
1	14,5 7,35	13,8 6,00	13,5 5,57	13,5 5,36	13,4 5,23	13,4 5,15	13,4 5,09	13,4 5,05	13,4 5,01	13,4 4,99	13,3 4,74
2	8,87 12,23	8,12 9,31	7,93 8,32	7,85 7,83	7,80 7,53	7,78 7,33	7,76 7,18	7,74 7,07	7,74 6,99	7,73 6,92	7,70 6,30
3	6,82 17,46	6,13 12,86	5,93 11,26	5,84 10,43	5,79 9,92	5,76 9,58	5,74 9,33	5,73 9,14	5,71 8,99	5,71 8,87	5,67 7,75
4	5,75 22,93	5,11 16,60	4,91 14,35	4,82 13,17	4,77 12,43	4,74 11,92	4,72 11,55	4,70 11,27	4,69 11,05	4,69 10,87	4,65 9,15
5	5,11 28,59	4,49 20,51	4,30 17,58	4,21 16,00	4,16 15,03	4,13 14,35	4,10 13,85	4,09 13,40	4,08 13,17	4,07 12,92	4,02 10,51
6	4,62 34,43	4,07 24,56	3,89 20,94	3,80 18,97	3,74 17,73	3,71 16,86	3,69 16,23	3,67 15,74	3,66 15,34	3,65 15,02	3,60 11,84
7	4,35 40,54	3,78 28,74	3,59 24,39	3,50 22,02	3,45 20,51	3,42 19,46	3,39 18,67	3,38 18,06	3,36 17,57	3,36 17,18	3,30 13,15
8	4,09 46,52	3,55 33,03	3,37 27,95	3,28 26,16	3,22 23,37	3,19 22,11	3,17 21,18	3,15 20,45	3,14 19,87	3,13 19,39	3,07 14,43
9	3,90 52,75	3,37 37,42	3,19 31,59	3,10 28,41	3,05 26,31	3,02 24,84	2,99 23,75	2,98 22,89	2,96 22,21	2,95 21,64	2,89 15,71
10	3,75 59,08	3,23 41,90	3,05 35,33	2,96 31,67	2,91 29,31	2,87 27,63	2,85 26,38	2,83 25,39	2,82 24,60	2,81 23,95	2,75 16,96

* *Bemerkung.* In jedem Quadrat: oben $k(n, c)$, unten $\eta(n, c)$.

4. Es sei $k_1 = 5$, $k_0 = 1$, also $k_1/k_0 = 5$. Die Ungleichung (10) wird durch die Schemata mit $c \geqq 4$, $n \geqq \max\{7, c + 1\}$ erfüllt,

Wir setzen voraus, daß $C(n, y) = 7{,}5y + n$. Man hat aus (11)

$$C(n, y) = 7{,}5\, k_1^{-1}\eta(n, c) + n = 1{,}5\eta(n, c) + n .$$

Tabelle III enthält einige Werte von $C(n, y)$.

TABELLE III

	$c = 4$	$c = 5$
$n =$ 7	28,52	37,76
8	27,76	34,37
9	27,64	33,00
10	27,88	32,54
11	28,32	32,52

Der Stichprobenplan (9,4) ist also optimal.

Wenn wir nur den Materialverbrauch minimisieren wollen, d. h. $C(n, y) = y$, die Länge x der Drahtabschnitte, deren Prüfung praktisch möglich ist, aber nach unten durch $a = 1$ beschränkt ist, so ergibt sich (8,9) als optimaler Stichprobenplan. Ist $a = 0{,}5$, so muß man (7,5) wählen.

ANWENDUNGSMÖGLICHKEITEN DER SIMULATIONSTECHNIK IM TEXTILMASCHINENBAU UND IN DER TEXTILINDUSTRIE

von

U. Liebscher

Technische Hochschule
Karl-Marx-Stadt, Deutsche Demokratische Republik

Die klassische Form des Unterrichts verlangt vom Lehrendem und vom Schüler aktive Teilnahme. Die Schwerpunkte der Aktivität liegen jedoch eindeutig beim Lehrer, während die Schüler eine gewisse Passivität aufweisen. Es wurden in den letzten Jahren und Jahrzehnten fortschrittliche Lehrmethoden entwickelt, die nicht auf dem Auswendiglernen allein basieren, sondern wo die sofortige praktische Anwendung ein wichtiges Schulungselement ist. Sie werden auch als aktive Lehrmethoden bezeichnet. Es müssen zum Beispiel in der Sowjetunion Studenten der Fremdsprachen Vorträge simultan übersetzen. Hierzu werden Konferenzen simuliert.

Aber auch auf anderen Gebieten wird bereits mit Simulierungen gearbeitet. Es werden Flugmodelle im Windkanal erprobt und bei der Konstruktion von Brücken und Staudämmen werden Belastungen und Strömungsverhältnisse an Modellen simuliert.

Dies sind sogenannte gegenständliche Modelle. Der Ingenieur will die positiven Eigenschaften seiner Schöpfung kennenlernen und lässt daher ein solches gegenständliches Modell von ihr anfertigen. Am Abbild seiner projektierten Wirklichkeit kann er die Eigenschaften studieren, die ihm wichtig erscheinen. Das Modell weist hierbei nur die für den Versuch relevanten Eigenschaften auf, zum Beispiel Form oder Masseverhältnisse. Alle unwesentlichen Eigenschaften werden weggelassen oder vereinfacht. Burger definiert: Gegenständliche Modelle bilden unter Abstraktion die Wirklichkeit nach.

Wenn weiter abstrahiert wird, erhält man sogenannte gedenkliche Modelle. Sie lassen sich in einer besonderen Symbolsprache darstellen, z. B. als mathematische Modelle. Man beschränkt sich auf den logischen Zusammenhang der Wirklichkeit. Auch die Planspiele, über die ich noch berichten werde, sind solche Denkmodelle.

Bei verschiedenen Ausbildungen ist es unbedingt nötig zu lernen, wie kritische Situationen rasch und richtig gemeistert werden. Die Reaktion muss in Fleisch und Blut übergehen, gewissermaßen als bedingter Reflex erfolgen. Es werden für Flugzeugführer Flugsimulatoren entwickelt, die Gewitter, Notlandungen oder Ausfall von Motoren simulieren. Die getroffenen Entscheidungen am Simulator oder ihre Auswirkungen erfordern neue Entscheidungen. Auch die Kosmonauten und Astronauten sind an solchen Simulatoren geschult worden.

Die Offiziere aller Nationen werden durch simulierte Gefechte oder sogenannte militärische Planspiele ausgebildet. Vor 20 Jahren wurden diese Sandkastenspiele der Stabsoffiziere auf eine wissenschaftliche Grundlage gestellt. Dies erfolgte besonders durch v. Neumann und Morgenstern und ihre Spieltheorie. Man erkannte bald, dass diese Grundlagen auch für die Leitung in Industrie und Wirtschaft große Bedeutung erlangen könnten.

Reine Glücksspiele hängen ausschließlich vom Zufall ab. Der Ausgang strategischer Spiele ist jedoch auch noch den Entscheidungen der Spieler während des Spielverlaufs unterworfen. Bei vielen Kartenspielen wird der Zufall durch die Kartenverteilung gewahrt. Außerdem üben die Spieler auf Grund ihrer Erfahrungen großen Einfluß auf den Ausgang aus. Vor Spielbeginn legen die Spieler bewußt oder unbewußt fest, wie sie versuchen wollen, das Spiel zu gewinnen (z. B. Grand-Spiel, mit Aufnahme des Skats). Beim Festlegen der Spielweise werden neue Entscheidungen aus dem Blickwinkel der einzelnen Spieler nötig. Auf Grund der Regeln sind früher getroffene Entscheidungen der Gegner — zum Beispiel Wert der Skatablage — oder Zufallseinflüsse — Karten anderer Spieler — nicht bekannt. Aus den zur Verfügung stehenden Informationen müssen die Spieler nun Entscheidungen über ihr Verhalten treffen. Sie müssen sich für das Spiel eine Strategie zurechtlegen. Nach Burger versteht man unter der Strategie eines Spielers einen Plan, der das Verhalten des Spielers und somit die in der jeweiligen Situation vom Spieler zu treffenden Entscheidungen festlegt. Ein erfahrener, kluger Spieler wird eine gute Strategie und ein schlechter Spieler eine mangelhafte wählen. In vielen Fällen wird es eine Strategie geben, die am schnellsten und mit geringsten Aufwand zum Ziel führt. Das ist die optimale Strategie.

Es entstanden ab 1956 Planspiele für betriebliche Entscheidungen. Daß die Planspiele aus militärischem Bereich gekommen sind, sollte nicht stören. Als vor 150 Jahren Gneisenau in den Befreiungskriegen das Heer reformierte, führte er eine strenge Trennung zwischen Stab und Linie ein, also zwischen dem planenden und beratenden sowie dem ausführenden Organ. Dieses Leitungsprinzip hat sich heute mit gewissen Modifikationen auch im zivilen Bereich zur Leitung und Lenkung von Staaten durchgesetzt. Die Plankommission ist der Stab, der Volkswirtschaftsrat die Linie. Es ist an der Zeit, daß sich dieses Leitungsprinzip noch ausgeprägter als bisher zur Leitung der Betriebe auch bei uns durchsetzt, wie es im Ausland größtenteils schon der Fall ist.

1956 entwickelte die American Management Ass. erstmalig ein Simulationsmodell für Entscheidungen der obersten Betriebsleitung. Heute sind schätzungsweise über 150 wirtschaftliche Spielmodelle im Einsatz. In den Jahren 1956 bis 1960 beteiligten sich etwa 30 000 Personen an Spielen. Die Dauer der Spiele ist verschieden. Sie liegt etwa zwischen 5 und 90 Stunden.

Viele Spiele benötigen elektronische Rechenanlagen. Als Nachteil der Planspiele wird häufig angeführt, daß die Durchführung sehr langwierig und kostspielig sei, besonders wenn elektronische Rechenanlagen verwendet werden müssen. In der letzten Zeit ist daher eine Abkehr zur manuellen Datenverarbeitung zu verzeichnen. Es wird mit Hilfe von Nomogrammen,

Tabellen oder bei unserem Spiel mit einem von uns entwickelten Maßwertsimulator gearbeitet.

Nach Kant ist eines der Hauptkriterien für ein Spiel die Freiheit von Zwecken im Gegensatz zur Arbeit. Dieses Moment der Freiheit gilt aber für die Planspiele nicht mehr, da sie mit ernster Absicht unternommen werden. Unterhaltungsspiele werden dagegen zur Zerstreuung und Unterhaltung durchgeführt. Die Planspiele haben sich als eigenes Gebiet entwickelt und sind heute nicht mehr alle mit spieltheoretischen Betrachtungen zu erfassen.

Die Grundidee der Planspieltechnik ist, das Betriebsgeschehen zu einem Modell zu reduzieren, also die Wirklichkeit im Modell zu simulieren. Bei betrieblichen Planspielen handelt es sich um Modelle, die bestimmte Entscheidungsbereiche des Betriebes auswählen und Umweltbedingungen künstlich erzeugen. Diese Modelle werden nicht gelöst, so wie man eine Dreisatzaufgabe löst, sondern sie werden durch die Spieler betrieben, indem diese laufend Entscheidungen treffen müssen. Als Nachteil wird angegeben, dass diese Modelle unrealistischer seien, weil eine Reihe praktischer Gegebenheiten abstrahiert wird. Gute und schlechte Modelle unterscheiden sich darin, ob die Auswahl des Wesentlichen gelungen ist oder nicht. Bei einem Wirtschaftsmodell sind zum Beispiel 500 verschiedene Parameter vorhanden. Vom pädagogischen Standpunkt aus ist die Einfachheit der Spielmodelle kein Nachteil, da der Überblick erleichtert wird.

Zwischen dem Simulationsmodell, das durch den Schiedsrichter oder ein Rechengerät betrieben wird, und den Spielern herrscht ein Informationsfluß. Einerseits werden Informationen über getroffene Entscheidungen zum Schiedsrichter gelangen müssen. Auf der anderen Seite fließen Informationen über die Wirkungen dieser Entscheidungen zu den Spielern. Die Modelle befassen sich nahezu ausschließlich mit anonymen Betrieben und Produktionen. Nur einige wenige Spiele haben konkrete Probleme zum Gegenstand.

Als Komponenten des Spiels sind Spielregeln, die Schiedsrichtergruppe und schließlich die Teilnehmer ebenfalls in Gruppen zusammengefasst vorhanden. Die Spielregeln umreißen den Aktions- und Reaktionsbereich. In Aktionsbereich werden menschliche Entscheidungen befasst und im Reaktionsbereich zu Ergebnissen verarbeitet. Aktion und Reaktion wirken als Ursache und Wirkung so zusammen, daß sich eine Ähnlichkeit mit einem Regelkreis ergibt.

Diese Planspiele zum Schulen betrieblicher Leitungskräfte werden im Ausland immer stärker verbreitet. Im Rahmen der Hochschulausbildung werden Planspiele seit etwa 1959 in der DBR sowie in der DDR eingesetzt, z. B. an den Universitäten Hamburg, Frankfurt, Giessen, an der TH Darmstadt und schließlich auch an der Technischen Hochschule in Karl-Marx-Stadt. Das an der Technischen Hochschule in Karl-Marx-Stadt entwickelte Spiel trägt dazu bei, neuartige Lehrmethoden wie sie gefordert werden, zu entwickeln und einzuführen.

Es gibt Spiele für die verschiedensten Gebiete, so z. B. für die Leitung von Betrieben, für den Busverkehr in Großhäfen, für Flugplätze, aber auch für Materialfluß- und Lagerhaltungsprobleme sowie für die »Ferti-

gungs steuerung«. Diese befassen sich mit Problemen der Serienfertigung und zeigen den Zusammenhang von Arbeitsvorbereitung, Beschaffung und Vertrieb.

Bei der Betrachtung der verschiedenen Spiele können Unterscheidungen getroffen werden. Es gibt Spiele, die sich mit der Einordnung des Betriebes in die gesamte Wirtschaft befassen. Es konkurrieren Betriebe miteinander. Diese Spiele sind für unser Wirtschaftssystem nur bedingt geeignet. Abwandlungen könnten zur Schulung unserer Außenhandelsorgane verwendet werden. Dies wäre zum Beispiel ein Spiel, das sich mit der Absatzstrategie befasst, also mit Vertretereinsätzen, Werbeetat und Werbemittel.

Planspiele zur Leitung von Betrieben können bei uns nicht vom Standpunkt aus durchgeführt werden, Konkurrenzbetriebe zu schlagen und den Marktanteil zu vergrößern. Sie sollten vielmehr mit dem Ziel unternommen werden, optimale Strategien zur Leitung und Lenkung des Betriebes unter bestimmten gegebenen Bedingungen und natürlich unter Einhaltung gesetzlicher Vorschriften zu finden. Zur Bewertung der Spiele können Steigerung der Arbeitsproduktivität, Lohnfonds, Betriebegewinn und ähnliche für unsere Betriebe wichtige Kennwerte benutzt werden.

Die zweite Gruppe von Spielen befasst sich mit innerbetrieblichen Vorgängen. Es werden wie erwähnt, Planspiele für die Beschaffung und Lagerplanung erfolgreich eingesetzt. Sie sollen Entscheidungen auf dem Gebiet der Materialwirtschaft simulieren, von verschiedenen Institutkräfteeinsätze in der Fertigung entwickelt. Ein solcher Betrieb umfasst zum Beispiel Rohmateriallager, Fertigungsanlagen, Fertigwarenlager und eine Forschungs- und Entwicklungsabteilung. Es werden sieben verschiedene Arten von Rohmaterial zur Herstellung benötigt, die unter Beachtung der Lieferzeiten zu bestellen sind. Das Simulationsmodell verarbeitet unzureichende Lagerbestände an Rohstoffen, Aufwendungen für Reparaturen, Wertung, Verwendung von Überstunden zur Produktion, Produktivitätsschwankungen der Arbeitskräfte und ermittelt die Kosten der Produktion. Die Spieler entscheiden über Materialfluß, Einkauf von Rohmaterialen, sie führen Bestandsüberwachung und Fertigungssteuerung durch. Schließlich müssen sie die Auslieferung der Fertigerzeugnisse an auswärtige Zwischenlager regeln.

Bei vielen Spielen ist die Wirkung einer Entscheidung zeitlich verzögert (Investitionsentscheidungen). Forschungsvorhaben werden erst nach einigen Spielperioden wirksam, wobei eine Periode jeweils einem längeren Zeitabschnitt von 1/4 bis 1 Jahr entspricht. Es ist also auch eine langfristige Planung nötig.

Bisher gab es noch kein Spiel für die Steuerung der Qualität in der Fertigung. Anläßlich einer Vortragsreise vor über zwei Jahren nach Hannover tauchte bei einem Gespräch die Idee auf, ein solches Spiel zu entwickeln. Es wurde vom Referenten zusammen mit Herrn Ahorner, Hannover, der erste Typ eines Qualitätssteuerungsspiels entwickelt, das in der DDR und der DBR mit einer großen Anzahl von Gruppen durchgespielt wurde. Dieses Spiel ist ein isoliertes Modell, da die Ergebnisse der Gruppen, die parallel spielen, sich nicht wechselseitig beeinflussen. Der Gegensatz wäre ein Interaktionsmodell.

Bei den Übungsspielen werden zwei, drei oder mehr Gruppen zu zwei bis vier Teilnehmern gebildet, die in getrennten Räumen arbeiten. Bei unserem Übungsspiel Qualitätssteuerung ist die Fertigung eines Teiles oder einer Ware, z. B. textiler Natur, gegeben. Vom Schiedsrichter erhalten die Gruppen auf einem Formular ungefähre Angaben über die Funktionsgenauigkeit von vier Maschinen bzw. Maschinensätzen. Sie können wählen, auf welcher dieser vier Maschinen die Fertigung durchgeführt werden soll. Sie erhalten dann verarbeitet in Tabellen oder in unserem Gerät die derzeitige Einstellung der gewählten Maschine durch den Einrichter. Nach bestimmten Gesichtspunkten entnehmen die einzelnen Gruppen diesen Tabellen oder dem Gerät Meßproben. Die gewonnenen Werte werden nach bekannten Methoden verarbeitet, so z. B. in Kontrollkarten eingezeichnet oder es wird anhand einer Strichliste der Mittelwert errechnet.

Dann muß eine Entscheidung gefällt werden, z. B. die Fertigung soll in der bisherigen Weise weitergeführt werden oder die Fertigung ist auf einer anderen Maschine durchzuführen.

Dann beginnt die nächste Periode, wobei jede Periode einer Schicht entspricht. Es wird so lange gespielt, bis die Fertigung beherrscht und die gewünschte Qualität erzeugt wird.

Durch den Schiedsrichter erfolgt die Bewertung des Aufwandes, der zur Erreichung des Zieles nötig war. Jede Messung, jede Verstellung der Maschine und jedes Ausschußteil kosten Punkte. Auch die Fertigung auf genauer arbeitenden Maschinen ist teuerer. Dasjenige Kollektiv hat gewonnen, das die Fertigung in kürzester Zeit und bei geringstem Kostenaufwand reguliert und beherrscht. Zum Schluß findet eine Ergebnisbesprechung (Manöverkritik) statt. Hier legen die einzelnen Gruppen ihre Gedankengänge dar, warum sie diese und nicht jene Entscheidung gefällt haben. Sie setzen sich hierbei der Kritik der anderen Gruppen aus, wobei sich meistens eine lebhafte Diskussion ergibt.

Auch bei unserem Spiel sind die Kernphasen des Entscheidungsprozesses eingebaut, wie sie Bleicher definiert:

1. Erkennen des Problems
2. Beschaffen und Beurteilen der Daten
3. Herausarbeiten alternativer Lösungsmöglichkeiten für das Problem
4. Bewerten dieser Lösungsmöglichkeiten
5. Entscheidung für die geeignete Lösung

Die Teilnehmer lernen bei unserem Spiel, wie auch bei allen anderen Spielen, eine Reihe von Fähigkeiten, so Fähigkeiten Entscheidungen zu treffen, Fähigkeiten zum Abstrahieren, Organisieren und Auswählen von Informationen aus einer komplexen und diffusen Umwelt, die Fähigkeit Voraussagen zu machen und zu planen, die Fähigkeit Eigenschaften eines Spezialisten mit denen eines Generalisten zu verbinden und schließlich und vor allem die Fähigkeit, mit anderen Menschen erfolgreich zusammenarbeiten zu können.

Dadurch, daß die Spieler in Gruppen zusammengefaßt werden, wird die Arbeit im Kollektiv — Gruppenarbeit — gefördert. Innerhalb der Gruppen müssen Entscheidungen einstimmig gefaßt werden, d. h. müssen

unter Umständen einzelne Teilnehmer ihre Interessen denen der Gruppe unterordnen. Das Arbeiten im Kollektiv wird sich auch bei der betrieblichen Arbeit später vorteilhaft auswirken.

Die Spieler wissen zwar, daß das Spiel dem Lernen dient. Während des Spiels geht jedoch das Gefühl für das Lernen verloren. Sie lernen nicht zielstrebig und bewußt, sondern durch Pendeln zwischen falschen und richtigen Aktionen, also durch Pendeln zwischen Versuch, Irrtum, Korrektur und neuem Versuch. Wichtige Momente sind ferner das erprobende Anwenden, das Einüben und Ausüben.

Die Mitwirkung erzeugt so viel Begeisterung, daß das Lernen weitgehend unterschwellig vollzogen wird. Es ist einleuchtend, daß dies die leichteste Form des Lernens ist. Bei allen Spielen konnte auch eine starke gefühlsmäßige Anteilnahme beobachtet werden.

Das Übungsspiel hat noch mehr Ziele. Es dient wie erwähnt der Ausbildung, da sich die Teilnehmer neues Wissen aneignen. Ferner wird die Anwendung bereits vorhandenen Wissens geübt, so daß Fähigkeiten vermittelt werden und das Spiel auch dem Training dient. Die Spiele können aber auch in der Forschung eingesetzt werden, indem die Verhaltensweise von Fertigungs- und Kontrollmethoden geprüft wird. Es können durch das Planspiel Qualitätssteuerung auch Änderungen in der Maschinenbelegung, im Fertigungs- und Prüfverfahren an einem Modell wirklichkeitsgerecht durchgeführt werden. Dies gestattet, optimale Fertigungsabläufe und Prüfverfahren zu erarbeiten, ohne daß durch Versuche die Produktionskapazität belastet wird und unnötiger Ausschuß entsteht.

Das erste Modell, das gemeinsam mit dem RKW Hannover erarbeitet wurde, wurde von uns für die Bedürfnisse der einzelnen Industriezweige im letzten Jahr noch variiert. Wir haben nun Spieltypen mit und ohne Trend und Störquellen. Beispiel für ein Trend ist Stahlabnutzung und für Störquellen Lunkerstellen im Material oder Änderung der Klimabedingung bei textiler Verarbeitung. Ein Spieltyp schließlich ist für die Prüfung von Eigenschaften, also für die Prüfung nicht messbarer Merkmale eingerichtet. Diese Grundtypen werden durch verschiedene Schwierigkeitsgrade und Zahlenwerte modifiziert, indem die Belange der Metallfertigung oder der Spinnerei oder Weberei und Wirkerei stärker beachtet werden.

Wir benutzen als Datenlieferanten von uns erarbeitete Tabellen oder den gemeinsam mit unserem wissenschaftlichen Mitarbeiter Sarembe entwickelten Meßwertsimulator. Der Meßwertsimulator wird mit Meßwerten nach einer vorgegebenen Verteilung programmiert. Jede Übungsgruppe bedient ein Gerät, aus dem der Meßwert durch Knopfdruck ausgewählt und abgefragt wird. Der Schiedsrichter sitzt am Steuerpult und kann zu jeder gegebenen Zeit die eingegebene Meßwertverteilung verändern. Dies geschieht bei den Spielen nach dem von uns festgelegten Programm.

Dieses Spiel soll auf Arbeitsvorbereitung und Materialversorgung erweitert werden und der Ausbildung von leitenden Mitarbeitern im Fertigungsbereich und Erprobung von betrieblichen Ablaufsystemen dienen. Zur Herstellung sind drei Materialarten erforderlich in festgelegten unterschiedlichen Mengen, wobei normaler Ausschuß berücksichtigt wird. Störungselemente sind u. a. Verzögerungen bei der Anlieferung der Roh-

stoffe. Die drei herzustellenden Produkte durchlaufen drei Maschinensysteme. Ausfallzeiten, die durch Maschinenstörungen, Stromausfall oder Arbeitskräfteausfall bedingt sein können, werden durch den Schiedsrichter mitgeteilt. Wichtig ist die richtige Materialbestellung. Mangel führt zu Maschinenstillständen, Terminüberschreitungen sowie Konventionalstrafen. Überhöhte Bestände ergeben erhöhte Lagerkosten und außerdem müssen die gesetzlichen Bestimmungen beachtet werden. Überstunden müssen vom Schiedsrichter genehmigt werden, der auch die Lohnzuschläge mitteilt. Es müssen Durchlaufpläne ausgearbeitet werden und Entscheidungen in folgenden Bereichen gefällt werden:

(a) Forschungs- und Entwicklungsvorhaben
(b) Bestimmung und Veränderung der Kapazität
(c) Bestimmung und Veränderung der Beschäftigtenzahl
(d) Preisliche Gestaltung
(e) Investkreditnahme

Die vorhandenen Spieltypen sowie das in Vorbereitung befindliche Spiel wird auf Anfrage mit Gruppen aus der Industrie durchgespielt. Es ist zu hoffen, daß die Übungsspiele dazu beitragen, das vermittelte Wissen im wahrsten Sinne des Wortes spielend zu erlernen.

ASYMPTOTICALLY OPTIMAL TESTS OF COMPOSITE HYPOTHESES FOR RANDOMIZED EXPERIMENTS WITH NON-CONTROLLED PREDICTOR VARIABLES

by

J. NEYMAN and E. L. SCOTT
University of California
Berkeley, U.S.A

1. INTRODUCTION

We begin by describing a particular problem which motivated the general theoretical developments given in the subsequent sections of the paper. Consider a randomized cloud seeding experiment conducted as follows.

There is a "target area" where it is intended to increase precipitation by seeding clouds. In addition, there are s "control" areas, somewhat isolated from the target so that cloud seeding over the target is assumed not to affect the precipitation in the controls. Not all storms passing over the given locality are suitable for seeding. When a storm is approaching, a meteorologist examines appropriate data and decides whether this storm offers an opportunity for seeding. In the affirmative case, a figurative coin is tossed and, if it falls heads, then the seeding operations are performed. If the coin falls tails, there is no seeding. In either case, observations are made on the precipitation both in the target and in the control areas. The effectiveness of seeding is judged from the comparison of the two sets of observations, one set referring to storms found seedable but not seeded and the other set to storms actually seeded.

The peculiarity of the situation described, which is not uncommon in other domains of experimentation, is that the observable random variables, in this case the amounts of precipitation in the control areas and in the target are far from normally distributed. In addition, the conditional variance of the amount of rain in the target, given the rain in the controls, depends very much on the values of the latter, which makes it inadvisable to use the familiar optimal tests of the normal theory. To obviate the difficulty, it is usual to apply a "norming" transformation of the directly observable variables and then to use tests of the normal theory, say the t-test. While, with a satisfactory norming transformation, this procedure maintains the intended level of significance, there is the difficulty that, while the normal test used may be optimal for transformed variables, this need not be the case with reference to alternative hypotheses expressed in terms of the original, non-transformed variables. This point will be described presently in some detail. For the present, we point out that the situation calls for the deduction of a test which is in some sense optimal. Such asymptotically optimal tests might be deduced using the theory published elsewhere [1]. However, an immediate application of this theory in the present case is complicated by the presence of predictor variables, namely of the amounts

of precipitation in the control areas the distribution of which is generally unknown.

The general problem is, then, as follows: to deduce a locally asymptotically optimal test of the hypothesis that the seeding of clouds has no effect on precipitation on the target using only specific hypotheses on the conditional distribution of the test variable (precipitation in the target) given the predictors (precipitation in the controls) and only very vague restrictions on the distribution of the predictors. The optimality of the test is desirable with respect to the set of alternatives expressed in terms of the non-transformed variables.

Proceeding to details, we denote by $X = (X_1, X_2, \ldots, X_s)$ the amounts of precipitation from the same seedable storm observable, respectively, in the s control areas. These amounts, representing predictor variables, may be measured in inches or be the results of some transformation which, however, is immaterial. The corresponding amount of precipitation in inches in the target is the "experimental" variable denoted by Y. We accept as given that the square root transformation of Y insures the linearity regression on X and, in fact, normality of conditional distribution given X, whether or not the storm is seeded and whether or not the seeding has an effect. In other words, following the traditional procedure, we set

$$Y = \eta^2, \tag{1}$$

where η is considered as a variable with conditional distribution given X represented by the density

$$p_{\eta \mid x}(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\left(y - a_0 - \sum_{i=1}^{s} a_i x_i\right)^2 \Big/ 2\sigma^2 \right\}, \tag{2}$$

where $\sigma, a_0, a_1, \ldots, a_s$ are unknown constants, possibly dependent on whether or not the given storm is seeded. From here on, the customary analysis of the experimental data consists in computing the value of $\eta = \sqrt{Y}$ for each storm and in using the normal theory tests of the hypothesis that the seeding had no effect. Quite frequently, this is done on the tentative hypothesis that the seeding does not affect the regression coefficients $a_1, a_2, \ldots, a_s$ but that it may affect only the constant term a_0. In this case the optimal normal theory test is the t-test.

Against this procedure the opinion of the knowledgeable meteorologists may be quoted that, if the seeding of clouds has an effect on precipitation in the target, this effect is multiplicative. In other words, if for a given X the expected non-seeded precipitation in inches is $Y(x)$, then the expected precipitation affected by seeding will be $Y(x)e^{2\xi}$ where ξ stands for the positive or negative conventional measure of the effect of seeding. Thus, in order to honor the expectations of the meteorologists, the test of the hypothesis of no effect of seeding should be made optimal with respect to the alternatives specifying a non-zero value of ξ.

Now let us consider what this implies with regard to the transformed variable η. Assuming that it is conditionally normal whether or not the storm is seeded and that its regression on X is linear, we write

$$\mathbf{E}(\eta \mid x, \xi = 0) = a_0 + \Sigma a_i x_i\,, \qquad \mathrm{Var}(\eta \mid x, \xi = 0) = \sigma^2 \tag{3}$$

for non-seeded storms, and

$$\mathbf{E}(\eta \mid x, \xi) = a_0(\xi) + \Sigma a_i(\xi) x_i\,, \qquad \mathrm{Var}(\eta \mid x, \xi) = \sigma^2(\xi) \tag{4}$$

for seeded storms, and using the assumption of multiplicativity of the effect of seeding, the assumption referring to the non-transformed variable $Y = \eta^2$, we obtain the identity

$$\begin{aligned} \mathbf{E}(Y \mid x, \xi) &= [a_0(\xi) + \Sigma a_i(\xi) x_i]^2 + \sigma^2(\xi) \equiv \\ &\equiv \{[a_0 + \Sigma a_i x_i]^2 + \sigma^2\} e^{2\xi} \equiv \mathbf{E}(Y \mid x, \xi = 0) e^{2\xi} \end{aligned} \tag{5}$$

for all x. Since both sides are polynomials of the second order in x, the identity requires that the coefficients of like powers of the x be equal. This in turn implies that

$$\sigma(\xi) = \sigma e^{\xi} \quad \text{and} \quad a_i(\xi) = a_i e^{\xi}\,, \quad i = 0, 1, 2, \ldots, s. \tag{6}$$

Thus, the optimality of the test of no effect of cloud seeding (or of any other treatment that may be used in other experiment) is required with respect to the alternatives represented by (6). Even though the underlying distribution of η (for a given x) is supposed normal with a variance independent of x, no test based on normal theory seems to be known that has optimal properties with respect to the alternatives (6). The construction of an asymptotically optimal test corresponding to this situation is hampered by the presence of the predictor variables x about whose distribution one is reluctant to adopt any hypotheses other than of a very general character such as, for example, that there is an underlying probability density depending on some parameters, that this density is several times differentiable under the integral sign and that certain moments exist.

The problem treated below represents a generalization of the particular problem just described. The treatment is based on theory given in [1] to which the reader is referred for various details.

2. GENERAL PROBLEM

CASE I. UNRESTRICTED RANDOMIZATION

We consider a randomized experiment with a single treatment performed on certain experimental units denoted by u (seedable storms, similarly diagnosed patients in a clinic, batches of raw material in a factory, etc.). Each experimental unit is characterized by a predictor variable X, possibly a vector. The distribution of X has a probability density $f(x|\Theta)$ depending upon a vector parameter $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_r)$ which satisfies the regularity

conditions specified in [1], including differentiability with respect to each Θ_i under the sign of the integral taken over the whole sample space of X. For each experimental unit a random experiment E is performed, independent of the predictor x and of similar experiments performed for other units. The probability of success in each experiment E is the same π, $0 < \pi < 1$. If the experiment E yields success, the corresponding experimental unit u is subject to a specified treatment. Otherwise, the unit is left untreated. In either case one observes on u a random variable Y, a scalar, which is the experimental random variable. It is assumed that for each possible value of the predictor X, the experimental variable Y has a known probability density, say $p(y|x, \vartheta)$ depending on a vector parameter $\vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_s) \in \Theta$ where Θ is an open set. This assumption is maintained whether or not the given experimental unit is subjected to treatment. However, there is an additional assumption regarding the effect of treatment conventionally measured by a parameter ξ capable of assuming values in an open interval Ξ containing zero. Namely, if without treatment the conditional distribution of Y given x corresponds to a parameter point $\vartheta \in \Theta$, then after treatment with effectiveness ξ the conditional distribution of Y given the same x corresponds to the parameter point $\vartheta(\xi)$ with

$$\vartheta_j(\xi) = \vartheta_i + \vartheta'_i\xi + 0(\xi), \qquad j = 1,2 \ldots, s, \tag{7}$$

where ϑ'_i is supposed to be a known function of ϑ. In the example described in the Introduction, the parameter vector $\vartheta = (a_0, a_1, \ldots, a_s, \sigma)$, with the a real and σ positive. Formulas (7) correspond to (6). Let the variable $t = 1$ if the experiment E performed for a given experimental unit gives success, and $t = 0$ otherwise. With this convention, the conditional density of Y given $X = x$ can be written as $p[y|x, \vartheta(\xi t)]$, whether or not the corresponding experimental unit is "treated".

In the above setup, the observable random variable corresponding to any given experimental unit is the triplet, say

$$\mathscr{X} = (X, t, Y) \tag{8}$$

of which X is a vector of dimensionality s. The sample space, say W, of $\mathscr{X}$ is a cartesian product of the sample space of X, of the sample space of t which consists of two points only 0 and 1, and of the sample space of Y. The probability density of $\mathscr{X}$ defined on W is, say,

$$\Psi(x, t, y) = \pi^t(1 - \pi)^{1-t} f(x \mid \theta) p\big[y \mid x, \vartheta(\xi t)\big]. \tag{9}$$

Now we assume that the consecutive experimental units $u_1, u_2, \ldots, u_n, \ldots$ are mutually independent so that the corresponding random vectors (8), that is, $\mathscr{X}_1, \mathscr{X}_2, \ldots, \mathscr{X}_n, \ldots$ are mutually independent. Since all of them have the same distribution characterized by (9), the problem of an asymptotically locally optimal test of the hypothesis that the treatment has no effect, that is, that $\xi = 0$, fits exactly the framework of the theory given in [1], subject only to the conditions of regularity specified therein. Assuming that these conditions are satisfied by both f and p, and assuming further

that for each of the parameters Θ_i and ϑ_j there is available an estimator $\hat{\Theta}_{in}$ and $\hat{\vartheta}_{jn}$ which is at least locally root n consistent, the family $C(\alpha)$ of asymptotic tests corresponding to the chosen level of significance α is defined as follows.

3. TESTS OF CLASS $C(\alpha)$

We choose a function g defined on W which has the property of being a Cramér function (see [1]). We set $\xi = 0$ and let $g(\Theta, \vartheta)$ and $\sigma^2(\Theta, \vartheta)$ be the expectation and variance of $g(\chi)$, respectively. Next we set

$$Z_n(g, \theta, \vartheta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{g(\mathscr{X}_i) - g(\theta, \vartheta)}{\sigma(\theta, \vartheta)}. \tag{10}$$

Obviously if $\xi = 0$, the distribution of Z_n tends to be normal $N(0, 1)$ as $n \to \infty$. Let $S(\alpha)$ be a Borel set on the real line subject to two conditions: (a) the characteristic function of $S(\alpha)$ is Riemann integrable and (b)

$$\frac{1}{\sqrt{2\pi}} \int_{S(\alpha)} e^{-x^2/2} dx = \alpha . \tag{11}$$

These conditions imply that

$$\lim_{n \to \infty} \mathbf{P}\{Z_n(g, \theta, \vartheta) \in S(\alpha)\} = \alpha . \tag{12}$$

If $Z_n(g, \Theta, \vartheta)$ could be calculated from the observations $\{\mathscr{X}_n\}$, then an asymptotic test of the hypothesis H asserting $\xi = 0$ could be obtained by making a rule of rejecting H whenever the calculated $Z_n(g, \Theta, \vartheta) \in S(\alpha)$. However, the calculation of $Z_n(g, \Theta, \vartheta)$ requires the knowledge of the parameters Θ and ϑ. Thus, instead of the exact values of those parameters we substitute in Z_n the estimators $\hat{\Theta}$ and $\hat{\vartheta}$ with the hope that the resulting random variable $Z_n(g, \hat{\Theta}, \hat{\vartheta})$ will also be asymptotically normal. As shown in [1], the condition necessary and sufficient for $Z_n(g, \hat{\Theta}, \hat{\vartheta})$ to differ from $Z_n(g, \Theta, \vartheta)$ by a random variable that tends to zero in probability [which implies the asymptotic normality of $Z_n(g, \hat{\Theta}, \hat{\vartheta})$] is that the function g be orthogonal to the partial derivatives of $\log \psi$ with respect to the nuisance parameters Θ and ϑ. Subjecting g to this last condition, a member of the class $C(\alpha)$ of tests is identified by the choice of the function g and of the set $S(\alpha)$. It is shown in [1] that the optimal test of this class is the one corresponding to, say,

$$g^* = \left[\frac{\partial \log \Psi}{\partial \xi} - \sum_{i=1}^{r} a_i \frac{\partial \log \Psi}{\partial \theta_i} - \sum_{j=1}^{s} b_j \frac{\partial \log \Psi}{\partial \vartheta_j} \right]_{\xi=0}, \tag{13}$$

where the coefficients a_i and b_j minimize the variance of (13), computed on the assumption that $\xi = 0$. Also, the set $S(\alpha)$ maximizing the asymptotic

power of the test based on (12) extends from a constant $v(\alpha)$ defined by (10) to infinity. The asymptotic power of the optimal test so obtained is then, say,

$$\beta_n(\xi) = \frac{1}{\sqrt{2\pi}} \int_{v(\alpha)}^{\infty} \exp\left\{-\frac{1}{2}\left[x - \xi\sqrt{n}\sigma^{*2}\right]^2\right\} dx, \tag{14}$$

where σ^{*2} stands for the minimum variance of (12) corresponding to $\xi = 0$.

4. LOCALLY ASYMPTOTICALLY OPTIMAL TEST OF CLASS $C(\alpha)$ FOR THE RANDOMIZED EXPERIMENT CONSIDERED

In order to obtain the optimal test of class $C(\alpha)$ for the situation described in Section 2 it is now sufficient to use the density (9) in order to perform the operations explained in Section 3. Let f_i denote the logarithmic derivative of f with respect to Θ_i and let $\varphi_j, j = 1, 2, \ldots, s$, denote the logarithmic derivatives of $p(y|x, \vartheta)$ with respect to ϑ_j. We have from (9)

$$\log \Psi = t \log \pi + (1 - t) \log (1 - \pi) + \log f(x \mid \theta) + \log p[y \mid x, \vartheta(\xi t)] \tag{15}$$

and

$$\left.\frac{\partial \log \Psi}{\partial \xi}\right|_{\xi=0} = t \sum_{j=1}^{s} \varphi_j(y, x, \vartheta)\, \vartheta_j', \tag{16}$$

$$\frac{\partial \log \Psi}{\partial \theta_i} = f_i(x, \theta), \qquad i = 1, 2, \ldots, r, \tag{17}$$

$$\left.\frac{\partial \log \Psi}{\partial \vartheta_j}\right|_{\xi=0} = \vartheta_j(y, x, \vartheta), \quad j = 1, 2, \ldots, s. \tag{18}$$

Expression (13) can now be written as

$$g^* = \sum_{j=1}^{s}(t\vartheta_j' - b_j)\,\varphi_j - \sum_{i=1}^{r} a_i f_i = \tag{19}$$

$$= (t - \pi) \sum_{j=1}^{s} \vartheta_j' \varphi_j + \sum_{j=1}^{s} (\pi\vartheta_j' - b_j)\,\varphi_j - \sum_{i=1}^{r} a_i f_i = (t - \pi) \sum_{j=1}^{s} \vartheta_j' \varphi_j + \chi, \text{ say.}$$

If $\xi = 0$, then the variable t is independent of both X and Y. Therefore, with $\xi = 0$, the variance of (19) corresponding to arbitrary coefficients a_i and b_i is

$$\mathrm{Var}(g^*) = \mathrm{Var}(t - \pi)\, \mathrm{Var}\left[\sum_{j=1}^{s} \vartheta_j' \varphi_j\right] + 2\, \mathbf{E}(\chi)\, \mathbf{E}(t - \pi) + \mathrm{Var}(\chi) \tag{20}$$

or, since $E(t) = \pi$, $\mathrm{Var}(t - \pi) = \pi\,(1 - \pi)$,

$$\mathrm{Var}(g^*) = \pi(1 - \pi)\, \mathrm{Var}\left(\sum_{j=1}^{s} \vartheta_j' \varphi_j\right) + \mathrm{Var}(\chi). \tag{21}$$

It will be noticed that the first term in the right side of (21) is independent of the coefficients a_i and b_j. As to the second term, Var (χ), it cannot be negative. Also, inspection of (19) shows that, by a proper choice of the a_i and b_j, the variable χ, and therefore its variance, can be made equal to zero. Namely, it is sufficient to set $a_i = 0$ for $i = 1, 2, \ldots, r$ and $b_j = \pi \vartheta_j'$. Thus the optimal function g^* is

$$g^* = (t - \pi) \sum_{j=1}^{s} \vartheta_j' \varphi_j(y, x, \vartheta) \tag{22}$$

and its variance σ^{*2} is

$$\sigma^{*2} = \pi(1 - \pi) \operatorname{Var} \left(\sum_{j=1}^{s} \vartheta_j' \varphi_j \right). \tag{23}$$

This expression shows that, for the above construction to lead to the desired result it is necessary and sufficient that the probability density p and the function $\vartheta(\xi)$ of (7) satisfy the condition that, with a positive probability under $\xi = 0$,

$$\sum_{j=1}^{s} \vartheta_j' \varphi_j \neq 0 . \tag{24}$$

Should the contrary be true, then the theory in [1] indicates another construction of the optimal test.

Finally, the optimal criterion of the form $Z_n(g^*, \hat{\Theta}, \hat{\vartheta})$, say Z_n^* is given by

$$Z_n^* = \frac{1}{\sqrt{n}} \frac{\sum_{i=1}^{n} (t_i - \pi) \sum_{j=1}^{s} \vartheta_j' \varphi_j(y_i, x_i, \hat{\vartheta})}{\{\pi(1 - \pi) \operatorname{Var} [\sum_{j=1}^{s} \vartheta_j' \varphi_j(Y, X, \hat{\vartheta})]\}^{1/2}} . \tag{25}$$

The essential point of the result obtained is that the optimal test criterion (25) is independent of the distribution f of the predictor variables with one possible exception. This is that the variance (23), which appears in the denominator of (25), may depend on the expectations of some functions of X. The nature of this possibility is illustrated in the example treated in the next section.

5. OPTIMAL TEST FOR THE RANDOMIZED CLOUD-SEEDING EXPERIMENT

Here we return to the randomized cloud-seeding experiment described in the Introduction. To simplify the formulas we shall assume that there is just one control area which will not diminish the illustrative value of the example. According to the preceding section, the construction of the optimal test criterion (25) requires the consideration of the conditional probability density of Y given the predictor $X = x$ and of Formula (7) describing the expected effect of the treatment. With only one control area we have

$$p(y \mid x, \vartheta) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ - \frac{1}{2\sigma^2} (y - a_0 - a_1 x)^2 \right\}, \tag{26}$$

where the symbol $\vartheta = (a_0, a_1, \sigma)$. Formulas (7) reduce to (6),

$$a_0(\xi) = a_0 e^{\xi},\ a_1(\xi) = a_1 e^{\xi},\ \sigma(\xi) = \sigma e^{\xi}, \tag{27}$$

which imply

$$a_0' = a_0,\ a_1' = a_1,\ \sigma' = \sigma. \tag{28}$$

The logarithmic derivatives of (26) are

$$\varphi_{a_0} = (Y - a_0 - a_1 X)/\sigma^2, \tag{29}$$

$$\varphi_{a_1} = X(Y - a_0 - a_1 X)/\sigma^2, \tag{30}$$

$$\varphi_{\sigma} = \frac{1}{\sigma}\left[\left(\frac{Y - a_0 - a_1 X}{\sigma}\right)^2 - 1\right]. \tag{31}$$

Thus, if the experiment is randomized with $\pi = 1/2$, the optimal g^* of Formula (22) is

$$g^* = \left(t - \frac{1}{2}\right)\left\{\frac{a_0'}{\sigma}\,\frac{Y - a_0 - a_1 X}{\sigma} + \frac{a_1' X}{\sigma}\,\frac{Y - a_0 - a_1 X}{\sigma} + \right.$$

$$\left. + \frac{\sigma'}{\sigma}\left[\left(\frac{Y - a_0 - a_1 X}{\sigma}\right)^2 - 1\right]\right\}. \tag{32}$$

Here we could use the values (28) but, for purposes of some comparisons to be considered presently there is the advantage of retaining the general notation a_0', a_1' and σ'. The variance of (32) corresponding to $\xi = 0$ is simply the expectation of the square of g^*. The first factor is independent of the rest and the expectation of its square is 1/4. The unconditional expectation of the square of the second factor is most easily computed by first obtaining its conditional expectation, given $X = x$, which reduces to the expectation of

$$\left[\frac{a_0' + a_1' x}{\sigma}\,\omega + \frac{\sigma'}{\sigma}(\omega^2 - 1)\right]^2 \tag{33}$$

where

$$\omega = \frac{Y - a_0 - a_1 x}{\sigma}, \tag{34}$$

is a normal variable with zero mean and unit variance. The result is

$$\left(\frac{a_0' + a_1' x}{\sigma}\right)^2 + 2\left(\frac{\sigma'}{\sigma}\right)^2. \tag{35}$$

Replacing x by X in (35) and taking the expectation with respect to the variation of X, the expected value of the square of the second factor in (32) is found to be

$$\frac{(a_0' + a_1' \mathbf{E}X)^2 + a_1'^2 \sigma_X^2}{\sigma^2} + 2\left(\frac{\sigma'}{\sigma}\right)^2, \tag{36}$$

where $\mathbf{E}X$ and σ_X^2 are the mean and the variance of the predictor X. Finally, the variance of g^* is found to be

$$\sigma^{*2} = \frac{1}{4}\left\{\frac{(a_0' + a_1'\mathbf{E}X)^2 + a_1'^2\sigma_X^2}{\sigma^2} + 2\left(\frac{\sigma'}{\sigma}\right)^2\right\}. \tag{37}$$

Here, then, as mentioned at the end of the last section, the variance (37) of the optimal g^* depends upon two parameters in the unknown distribution of the predictor X, namely on $\mathbf{E}X$ and on the variance σ_X^2. Under only very mild assumptions regarding the distribution of X these two parameters can be consistently estimated by the sample mean $\bar{x}$ and the sample variance say S^2. It follows that the optimal test criterion Z_n^* is given by

$$Z_n^* = \frac{1}{\sqrt{n}} \frac{\sum_{i=1}^{n}\left(t_i - \frac{1}{2}\right)\left\{\frac{a_0' + a_1'x_i}{\hat{\sigma}}\frac{y_i - \hat{a}_0 - \hat{a}_1x_i}{\hat{\sigma}} + \frac{\sigma'}{\hat{\sigma}}\left[\left(\frac{y_i - \hat{a}_0 - \hat{a}_1x_i}{\hat{\sigma}}\right)^2 - 1\right]\right\}}{\frac{1}{2}\left\{\frac{(a_0' + a_1'\bar{x})^2 + a_1'^2S^2}{\hat{\sigma}^2} + 2\left(\frac{\sigma'}{\hat{\sigma}}\right)^2\right\}^{1/2}}. \tag{38}$$

where $\hat{a}_0$, $\hat{a}_1$ and $\hat{\sigma}^2$ are the usual least squares estimators

$$\begin{aligned} \hat{a}_1 &= \frac{\Sigma(x_i - \bar{x})y_i}{\Sigma(x_i - \bar{x})^2} \\ \hat{a}_0 &= \bar{y} - \hat{a}_1\bar{x} \\ \hat{\sigma}^2 &= \Sigma[y_i - \bar{y} - \hat{a}_1(x_i - \bar{x})]^2/(n-2)\,. \end{aligned} \tag{39}$$

The non-centrality parameter in Formula (13) for the asymptotic power function is, say,

$$\tau_1 = \frac{\sqrt{n}\xi}{2}\left\{\frac{(a_0' + a_1'\mathbf{E}X)^2 + a_1^2\sigma_X^2}{\sigma^2} + 2\left(\frac{\sigma'}{\sigma}\right)^2\right\}^{1/2}. \tag{40}$$

If the asymptotically optimal test is needed against the alternatives given by Formulas (6), then all that is required is to use (38) and replace a_0' by $\hat{a}_0$, a_1' by $\hat{a}_1$ and σ' by $\hat{\sigma}$. However, it is interesting to consider the situation where the alternatives are

$$a_0(\xi) = a_0 + \xi \tag{41}$$

with $a_1(\xi) = a_1$, $\sigma(\xi) = \sigma$. Then $a_0' = 1$ and $a_1' = \sigma' = 0$. The optimal g^* function reduces to, say,

$$g^{**} = \left(t - \frac{1}{2}\right)(Y - a_0 - a_1X)/\sigma^2 \tag{42}$$

and the corresponding optimal test criterion is, say,

$$Z_n^{**} = \sum_{i=1}^{n}(2t_i - 1)(y_k - \hat{a}_0 - \hat{a}_1x_i)/\hat{\sigma}\sqrt{n} \tag{43}$$

which is a modification, say a randomized modification, of the familiar t-test. The corresponding noncentrality parameter is, say,

$$\tau_2 = \sqrt{n}\,\xi/2\sigma\,. \tag{44}$$

Now consider the third situation in which the true alternatives are given by (6) but, in ignorance of this, the test criterion used is Z_n^{**}, given by (43). In this case, as shown in [1] the asymptotic power function will be of the same form (14), but with the non-centrality parameter, say τ_3 equal to τ_1 multiplied by the correlation coefficient, say ϱ, between the optimal g^* function and the corresponding g^* function underlying the criterion Z_n^{**} actually used. The square of this correlation coefficient will then be the measure of the relative efficiency of the not-quite appropriate test criterion Z_n^{**}, compared with the optimal test criterion Z_n^*. Also, if the actual alternative hypotheses correspond to (41) while the test used is Z_n^*, the same number ϱ^2 will measure the relative efficiency of the test used. This consideration indicates the importance of guessing correctly the distribution of the observable variables in cases where the hypothesis tested is not true. Easy calculations show that in the present case,

$$\varrho^2 = \frac{[a_0 + a_1 \mathbf{E}X]^2}{[a_0 + a_1 \mathbf{E}X]^2 + a_1^2\sigma_X^2 + 2\sigma^2} \tag{45}$$

and it is seen that, depending upon the values of the constants involved, the use of an inappropriate test criterion may result in a substantial loss in efficiency.

6. OPTIMAL TEST

CASE II. RANDOMIZATION OF PAIRS OF EXPERIMENTAL UNITS

In Section 4 we considered an unrestricted randomization of the experiment. In the present section we shall assume that the sequence of forthcoming experimental units is divided into pairs, say, (u_{1i}, u_{2i}) for $i = 1,2, \ldots, n, \ldots$, and that each such pair is randomized independently from the others. A coin with probability π of falling heads is tossed, independently from other such tosses and independently from the predictor variables (X_1, X_2). If the coin falls heads, then the first unit u_{1i} is subject to treatment but not u_{2i} and vice versa. In this case the predictor variables (X_1, X_2) corresponding to any given pair (u_1, u_2) need not be assumed independent and, in fact, it may be natural to assume that they are dependent. For example, this will be the case when the pairs of experimental units are pairs of twins. However, we shall maintain the assumption that the pairs of predictors (X_{1i}, X_{2i}) corresponding to consecutive pairs of experimental units are mutually independent and identically distributed with a density $f(x_1, x_2|\Theta)$ dependent upon some vector parameter Θ. Our additional hypothesis regarding f will be that it is symmetric with regard to x_1 and x_2, so that the marginal distributions of X_1 and X_2 are the same.

Our present hypotheses regarding the experimental variables (Y_1, Y_2) corresponding to a pair of experimental units will be the same as before, with the addition that, given $X_1 = x_1$ and $X_2 = x_2$, the variables Y_1 and Y_2 are mutually independent with probability densities

$$p[y_1 \mid x_1, \vartheta(\xi t)] \quad \text{and} \quad p\{y_2 \mid x_2, \vartheta[\xi(1-t)]\}, \tag{46}$$

respectively. Each pair (u_1, u_2) is now characterized by a quintuple of observable random variables $\mathscr{X} = (X_1, X_2, t, Y_1, Y_2)$ with joint probability density, which we again denote by ψ,

$$\Psi = \pi^t(1-\pi)^{1-t} f(x_1, x_2 \mid \theta) p[y_1 \mid x_1, \vartheta(\xi t)] p\{y_2 \mid x_2, \vartheta[(\xi(1-t)]\}. \tag{47}$$

From now on, the operations leading to the locally asymptotically optimal criterion to test the hypothesis that $\xi = 0$ are perfectly analogous to what was described in the earlier sections of the paper and we restrict ourselves to indicating the results.

With obvious notation, the optimal g^* in this case is

$$g^* = \left(t - \frac{1}{2}\right) \sum_{j=1}^{s} [\varphi_j(y_1, x_1, \vartheta) - \varphi_j(y_2, x_2, \vartheta)]\vartheta_j' \tag{48}$$

independent of the value of π. Its variance is

$$\sigma^{*2} = \frac{1}{4} \mathbf{E}\left[\sum_{j=1}^{s} [\varphi_j(Y_1, X_1, \vartheta) - \varphi_j(Y_2, X_2, \vartheta)]\vartheta_j'\right]^2 =$$

$$= \frac{1}{2} \mathbf{E}\left[\sum_{j=1}^{s} \varphi_j(Y, X, \vartheta)\vartheta_j'\right]^2. \tag{49}$$

Formula (49) results from the following considerations:

(a) since the marginal distributions of X_1 and X_2 coincide it follows that

$$\mathbf{E}[\sum_{j=1}^{s} \varphi_j(Y_1, X_1, \vartheta)\vartheta_j']^2 = \mathbf{E}[\sum_{j=1}^{s} \varphi_j(Y_2, X_2, \vartheta)\vartheta_j']^2 \tag{50}$$

and (b) given $X_1 = x_1$ and $X_2 = x_2$, the variables Y_1 and Y_2 are mutually independent and therefore

$$\mathbf{E}[\varphi_j(Y_1, x_1, \vartheta)\, \varphi_k(Y_2, x_2, \vartheta)] = 0 \tag{51}$$

for $j, k = 1, 2, \ldots$. Using (48) and (49) it is easy to construct the optimal criterion Z_n^*.†

REFERENCE

1. NEYMAN, J., Optimal Asymptotic Tests of Composite Hypotheses. *Probability and Statistics* (The Harald Cramér Volume), Almquist and Wiksells, Uppsala, Sweden, pp. 416–444, 1959.

† This paper, delivered at the Conference by E. L. Scott, represents a version of an article subsequently published in the *Journal of American Statistical Association*, **60**, 699—721 (1965).

ÜBER NICHTLINEARE OPTIMIERUNG

von

J. ODERFELD

Mathematisches Institut der Polnischen Akademie der Wissenschaften
Warschau, Polen

1. EINFÜHRUNG

Bekanntlich versteht man unter Optimierung (auch Programmierung genannt) die folgende Aufgabe. Es seien $F(\mathrm{x})$ und $f_j(\mathrm{x})$, $j = 1, 2, \ldots, m$, Funktionen des Vektors $\mathrm{x} = [x_1, \ldots, x_n]$. Mit R bezeichnen wir den Bereich des n-dimensionalen Raumes, in welchem $f_j(\mathrm{x}) \leq 0$ für alle j. Es bezeichnen: $F(\mathrm{x})$ die Zielfunktion, $f_j(\mathrm{x}) \leq 0$ die Nebenbedingungen, R den zulässigen Bereich. Man sucht das absolute Minimum der Zielfunktion im zuläßigen Bereiche, also $\min \{F(\mathrm{x}) \mid \mathrm{x} \in R\}$.

Falls $F(\mathrm{x})$ und alle $f_j(\mathrm{x})$ linear sind, nennt man die Optimierung ebenfalls linear. Die lineare Optimierung ist gegenwärtig ein weitgehend erforschtes Gebiet und es steht daher dem Praktiker eine große Auswahl von Lösungsverfahren zur Verfügung. Im Falle einer nichtlinearen Optimierung entstehen aber weit größere Schwierigkeiten. Nur einige Sonderfälle, wie quadratische und konvexe Optimierung, der Fall der differenzierbaren Funktionen usw. sind bisher gründlich untersucht worden. Diese Einschränkungen sind aber für den Praktiker unbequem, da er schon beim Formulieren der Aufgabe Vereinfachungen in den Annahmen machen muß.

Bei Benutzung der statistischen Lösungsmethoden, also bei den sogenannten Monte-Carlo Methoden, fallen solche Einschränkungen weg. Es soll in diesem Zusammenhang kurz die Grundkonzeption der statistischen Methoden in Erinnerung gebracht werden. Mit Hilfe einer automatischen Rechenmaschine bildet man eine Folge der zufälligen Punkte, die eine gegebene (meistens gleichmäßige) Verteilung in einem n-dimensionalen Würfel aufweisen. Für jeden Punkt prüft man, ob er in dem zuläßigen Bereich liegt, ob also alle Nebenbedingungen erfüllt sind. Wenn dies der Fall ist, berechnet man den Wert der Zielfunktion und vergleicht ihn mit den Werten in den früher untersuchten Punkten. Als Lösung wählt man denjenigen Punkt, in welchem die Zielfunktion den kleinsten Wert erreicht hat.

Selbstverständlich ist dies eine Lösung im stochastischen Sinne. Sie strebt gegen der wahren Lösung, wenn die Zahl der Punkte unbeschränkt wächst und außerdem, wenn einige zusätzliche Bedingungen eingehalten werden.

Nach dieser Einführung möchte ich einige Probleme der Monte-Carlo Optimierung besprechen, mit welchen ich und meine Mitarbeiter aus Warschau in der Praxis bereits zu tun hatten. Dabei werde ich weder auf strenge mathematische Beweise noch auf technische Einzelheiten eingehen.

2. ORGANISATION DER ZUFALLSPROBEN

Die vorherskizzierte Methode ist eigentlich ein automatisiertes Probieren. Es besteht aus Zufallsproben, deren Planung aber einen großen Einfluß auf die Resultate hat. Es sollen zwei Fälle näher betrachtet werden.

2.1. Nehmen wir an, daß die Zone des absoluten Minimums im voraus (also vor den Proben) ziemlich gut bekannt ist. Dies ist z. B. der Fall, wenn es sich darum handelt, einen Transportplan zu prüfen, der seit mehreren Jahren ausgeführt wird und ständig empirisch verbessert worden ist. Man kann dann vermuten, daß die Punkte x^0 und x^e, die dem absoluten Minimum F_0 und dem experimentalen Minimum F_e entsprechen, nicht weit voneinander entfernt sind.

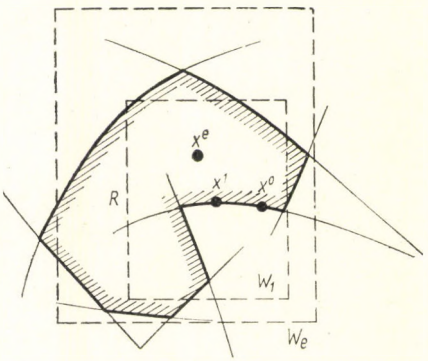


ABB. 1

Nun kann man die Abmessungen und die Lage des n-dimensionalen Würfels W_e derart wählen, daß der unbekannte Punkt x^0 mit praktischer Sicherheit umfaßt wird. Nach einem genügend langen und sinngemäßen Durchsuchen des Würfels W_e werden wir uns dem Punkte x^0 beliebig nähern. Man kann aber die Konvergenz beschleunigen, wenn man das Durchsuchen im Falle unterbricht, wenn ein derartiger Punkt x^1 gefunden wurde, in dem der Wert $F(x^1)$ erheblich kleiner als $F(x^e)$ ist. Rings um den Punkt x^1 baut man einen Würfel W_1, dessen Abmessungen kleiner als diejenigen des Würfels W_e sind und man wiederholt dieses Verfahren. Diese Konzeption wird im Abb. 1 illustriert.

Der Würfel W_e kann auch unbeschränkt gebildet werden. In einem solchen Fall wählt man die Verteilung der Zufallspunkte derart, daß die Dichte ihr Maximum im Punkte x^e erreicht.

2.2 Etwas komplizierter ist der Fall, wenn wir am Anfang des Verfahrens nur über geringe nicht-mathematische Information verfügen. Dies ist z. B. der Fall, wenn es sich darum handelt, die Hauptparameter einer völlig neuen technischen Anlage zu bestimmen. Besonders bei einer hohen Anzahl von Variablen ist es schwer zu erraten, wo das Minimum liegen könnte. Manchmal kann man sogar nicht sagen, ob der zulässige Bereich kohärent ist.

Abb. 2 illustriert einen solchen Fall. Der Bereich besteht hier aus zwei separaten Teilen R_1 und R_2, wobei R_1 viel kleiner als R_2 ist. Das absolute Minimum liegt im Punkt x^0 am Rande von R_1, was uns aber unbekannt ist. Deshalb sind wir gezwungen, den Würfel W_e aufs Geratewohl zu wählen. Dabei wäre es sogar möglich, den Teil R_1 des Bereichs unbewußt auszuschließen, was zu einer völlig falschen Lösung, z. B. im Punkt x^1 führen könnte.

Auch im unendlich großen Würfel W_e und bei ungleichmäßiger Verteilung der Zufallspunkte ist es wohl möglich, daß infolge der geringen Dichte keiner der Punkte in R_1 fällt.

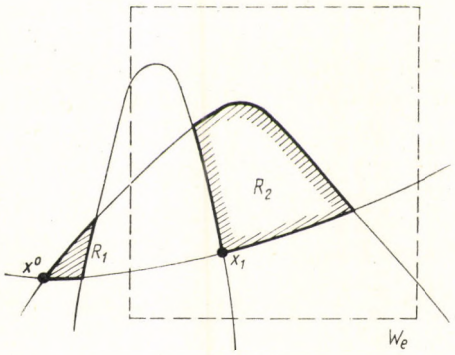


ABB. 2

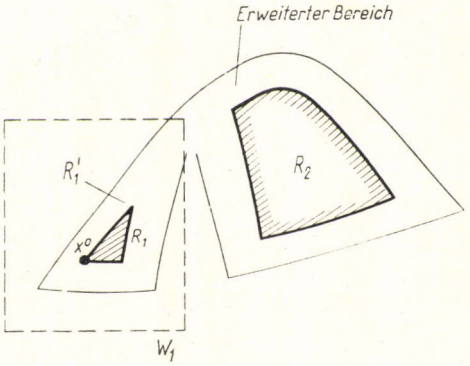


ABB. 3

Deshalb ist es zweckmäßig den zuläßigen Bereich vorübergehend zu erweitern, indem man alle oder einige Nebenbedingungen $f_j(x) \leq 0$ durch $f_j(x) \leq \varepsilon_j$ ($\varepsilon_j > 0$) ersetzt. Falls in einer Zone des erweiterten Bereichs erheblich kleinere Werte der Zielfunktion als in dem Originalbereich gefunden sind (z. B. in der Zone R_1' auf Abb. 3), umschließt man diese Zone mit einem Würfel W_1' und unterwirft sie einem eingehenden Durchsuchen unter den Bedingungen, die dem Originalbereich entsprechen.

Sämtliche hier erwähnte Schritte eignen sich zum Programmieren für automatische Rechenmaschinen.

Gegebenenfalls kann man auch andere Varianten des Verfahrens in Anspruch nehmen. In diesem Kurzbericht werden sie nicht besprochen.

3. GÜTE DER LÖSUNG

In diesem Kapitel beschränke ich mich auf das Durchsuchen in einem gegebenen Würfel.

Da der Zeit- und Kostenaufwand mit der Zahl der Zufallspunkte wächst, brauchen wir ein Kriterium, um das Durchsuchen beendigen zu können, sobald die Lösung genügend genau ist. Ein solches Kriterium kann auf verschiedene Weise erstellt werden.

3.1. Vor dem Durchsuchen bestimmt man die Anzahl N_1 der Punkte. die in den zulässigen Bereich fallen sollen. Nachdem die Zahl N_1 erreicht ist, nimmt man den Punkt des empirischen Minimums als Lösungspunkt.

Man kann aber gleichfalls die Anzahl N_2 der nacheinander folgenden Proben bestimmen, die zu keiner Verbesserung des Minimalwertes der Zielfunktion mehr führen.

Mit jedem der beiden Kriterien kann man eine Aussage verbinden, die die Güte der Lösung in der Sprache der mathematischen Statistik erfaßt. Auf diesem Wege hat R. Zieliński vom Mathematischen Institut der Polnischen Akademie der Wissenschaften interessante Resultate erhalten.

3.2. Man kann auch ein Extrapolationsverfahren anwenden. Bezeichnen wir mit $J(\varphi)$ die Anzahl der Punkte im zulässigen Bereiche, in welchem für die Zielfunktion $F \leq \varphi$ besteht. Die empirische Funktion $J(\varphi)$ ist

monoton. Wenn man sie bis zu $J = 0$ extrapoliert, bekommt man einen Wert $\hat{\varphi}$, der den Schätzwert für das absolute Minimum F_0 liefert. Die Differenz $F(x^e) - \hat{\varphi}$, wobei $F(x^e)$ den niedrigsten empirischen Wert bezeichnet, gibt einen direkten Hinweis auf die Güte der Lösung im Punkte x^e.

Zum Extrapolieren kann man verschiedene Methoden anwenden. Gut hat sich ein im folgenden beschriebenes Verfahren bewährt. Unter der Voraussetzung eines nicht allzu unregelmäßigen Charakters der Berandungen des Bereichs und der Zielfunktion ist es immer möglich ein derartiges Segment der empirischen Kurve $J(\varphi)$ zu finden (Abb. 4), daß

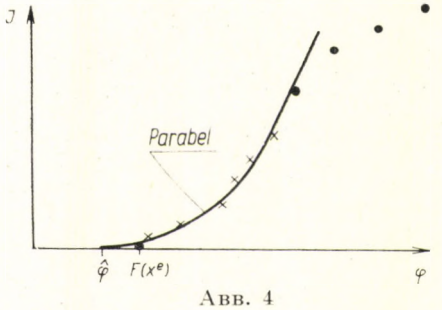


Abb. 4

1. die φ-Werte nicht viel größer als das beobachtete Minimum $F(x^e)$ sind;
2. die empirischen Punkte (φ, J) auf einen Verlauf ohne Wendepunkt der Kurve hinweisen.

Dieses Segment wird mit einer Parabel $J = C(\varphi - \hat{\varphi})^p$ approximiert. Die Parameter C, $\hat{\varphi}$, p (von denen eigentlich nur $\hat{\varphi}$ von Bedeutung ist) werden nach dem bekannten Verfahren der kleinsten Quadrate bestimmt, wobei man noch die Abhängigkeit der Streuung der Ordinaten der Approximationskurve von den empirischen J-Werten leicht berücksichtigen kann.

4. EINIGE ANWENDUNGEN

Die hier skizzierten Methoden sind in der Praxis erprobt worden. J. Golinski von der Polytechnischen Hochschule in Warschau hat sie mit Erfolg zur Konstruktion von Zahnradgetrieben und Brücken, sowie zur Lösung von Schwingungsproblemen angewendet.

In allen diesen Fällen reichte die angeführte Monte-Carlo Methode völlig aus und die geringen Abweichungen der gefundenen Lösungen von den wahren wurden durch die Einfachheit der angesetzten Methode ausgeglichen.

ON THE CONNECTEDNESS OF RANDOM GRAPHS

by

I. PALÁSTI

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary

The graph theory can be applied as a model in communication nets, in some chemical problems, in the study of networks, in econometry, in telecommunications, etc. The first question to arise was whether these graphs were connected or not. This is both mathematically fascinating and practically applicable. This is the reason why we deal with the connectedness of random graphs.

Although we shall not be concerned with questions of ordinary graphs, let us explain first what we mean by a graph. A graph consists of a finite non-empty set of labelled points $P_1, P_2, \ldots, P_n$ and a set of different pairs (P_i, P_j), where $(i \neq j)$. The points are called vertices, and the pairs the edges of the graph. We denote the number of edges by N.

We shall call random graphs those $\Gamma_{n,N}$'s which have n given labelled vertices and N edges chosen at random so that all the possible choices are supposed to be equiprobable.

P. Erdős and A. Rényi investigated in their paper [1] the connectedness of random graphs $\Gamma_{n,N}$ in the case of $n \to \infty$. They have shown that, if the number of edges is

$$N_c = \left[\frac{n}{2}\log n + cn\right]. \tag{1}$$

where c is an arbitrary, fixed real number ($[x]$ the integer part of the number x) then if we denote by $P_0(n, N_c)$ the probability that the graph $\Gamma_{n,N}$ is connected, then

$$\lim_{n\to\infty} \mathbf{P}_0(n, N_c) = e^{-e^{-2c}}. \tag{2}$$

They have raised the question, what would be true in the case of bichromatic graphs.

Paper [2] gives the answer but before going into it, let us consider the definitions. A graph is bichromatic if it has m given labelled points $P_1, P_2, \ldots, P_m$ of the first colour and n given labelled points $Q_1, Q_2, \ldots, Q_n$ of the second colour and N edges connecting points having different colours. We can imagine that all the points are coloured e.g. all the points P_i are red, but all Q_j are blue. The restriction is made that no vertices of the same colour are allowed to be connected by an edge. We obtain a bichromatic

random graph if we choose N different edges among the possible mn edges, supposing that all choices of the edges have the same probability. This probability is equal to $1/\binom{mn}{N}$. Let us denote this type of graph by $\Gamma_{m,n,N}$.

A bichromatic graph $G_{m,n,N}$ is connected, if any P_i can be reached from any Q_j by a path in $G_{m,n,N}$.

It was shown in the case of $m \sim \lambda n$ that the probability of the random graph $\Gamma_{m,n,N_{c,\lambda}}$ being connected, tends to the limit $e^{-e^{-c}}$, that is

$$\lim_{n\to\infty} \mathbf{P}(m, n, N_{c,\lambda}) = e^{-e^{-c}}, \tag{3}$$

if the number of the chosen edges is

$$N_{c,\lambda} = [m \log m, cm]. \tag{4}$$

In the special case of $\lambda = 1$, the limit of the probability of the connectedness of bichromatic random graph is $e^{-2e^{-c}}$, if $N_c = [n \log n + cn]$.

The connectedness of directed random graphs will now be considered.

An edge is called directed, if the occurrence order of its vertices is also given, that is, if we say that P_i is the initial point, and P_j is the terminal vertex of the edge. A directed edge is an outgoing edge with respect to its initial vertex, and an incoming edge, with respect to its terminal vertex.

A graph is called directed, if all of its edges are directed.

A directed random graph $\vec{\Gamma}_{n,N_c}$ is obtained, if we choose N different edges among the possible n^2 (directed) edges connecting n given vertices so that each of the $\binom{n^2}{N}$ possible choices is equiprobable. (This means that we allow also the loops, that is, the edges (P_i, P_i), too.) We suppose furthermore that there are no parallel edges, with the same direction, connecting two points.

A directed graph is connected if every pair of its points is joined by a semi-path (i.e. by a sequence of disorderly directed edges). The definitions of various connectedness, in the case of ordinary directed graphs and the applications of it, can be found in the paper of B. Roy [3].

It can be seen easily from the above mentioned definition of connectedness that the limit of the probability of a directed random graph being connected is the same as the probability [2] given by P. Erdős and A. Rényi. There is only the difference that an arbitrary direction is given to the edges of the graph obtained, if we do not allow the loops and any kind of parallel edges.

Therefore in the case of directed random graphs investigation of the strong connectedness seems to be more interesting.

A directed graph is strongly connected if every pair of its points can mutually be reached by a directed path.

For such graphs the following is true:

THEOREM. *If* $\mathbf{P}(n, N_c)$ *denotes the probability that the directed random graph* $\vec{\Gamma}_{n,\,N_c}$ *is strongly connected, supposing that the number of the edges is*

$$N_c = [n \log n + cn] \tag{5}$$

then

$$\lim_{n\to\infty} \mathbf{P}(n, N_c) = \mathrm{e}^{-2\mathrm{e}^{-c}}. \tag{6}$$

Here we do not give a complete proof. It can be found in our paper, [4]. It should only be noted that the chain order of the ideas of the proof is similar to the ideas of the proof of the above mentioned theorem of P. Erdős and A. Rényi. The outline of our proof is the following:

Let us classify the directed random graphs in the following way:

Let us consider to be of type A each of those directed graphs which consists of a strongly connected subgraph and of such semiisolated points k_1 ($k_1 = 0, 1, \ldots$). into which may lead incoming edges, but where there are no outgoing edges (say *sinks*, let us denote by K^- the set of these kind of points), and of such semi-isolated points k_2 ($k_2 = 0, 1, \ldots$) which may have only outgoing edges, but where there are no incoming ones (say *sources*, let K^+ be the set of these points). (Of course the graphs may have perfectly isolated points too. These are considered as the elements of the intersection of the point set K^+, and K^-, that is to the elements of the set $K^+ \cap K^-$.) All the other graphs belong to the type $\bar{A}$.

And we have to use the following:

LEMMA. Let us denote by $P(\bar{A}, n, N_c)$ the probability of the directed random graph $\vec{\Gamma}_{n,N_c}$ belonging to the type $\bar{A}$. Then

$$\lim_{n\to\infty} \mathbf{P}(\bar{A}, n, N_c) = 0. \tag{7}$$

That is to say, if n is large enough, then "almost all" directed random graphs $\vec{\Gamma}_{n,N_c}$ belong to the type A.

Summarizing the results we have seen that the limit of the probability of the directed random graph being strongly connected, is equal to the limit of the probability of the bichromatic random graph $\Gamma_{n,\,n,\,N_c}$ having $P_1, P_2, \ldots, P_n$ red and $Q_1, Q_2, \ldots, Q_n$ blue points (i.e. those bichromatic random graphs in which the number of the elements of different coloured point sets are asymptotically equivalent) being connected, supposing that the number of the chosen edges is also the same, namely $N_c = [n \log n + cn]$. This result is not so surprising because it is known in the graph theory that there exists a mapping (by simple cutting into two) of the vertices of directed (ordinary) graphs carrying over these graphs into bichromatic ones. But this result is interesting because this mapping does not preserve the connectivity of the graphs.

As an application of the last theorem, i.e. the strong connectedness of directed random graphs one can imagine the following:

There are e.g. n observatories. And it is very important that every observatory station be connected with any other one (by phone, or radio,

etc.) in both directions. Then if we wanted to control whether these connections are operating or not, then we need not measure, or control each of the possible n^2 connections, but after the $n(\log n + c)$ controling measurements we can say that the whole equipment is operating with probability $e^{-2e^{-c}}$ if the observatory pairs for measurement have been chosen at random.

REFERENCES

1. Erdős, P., Rényi, A., On Random Graphs I. *Publ. Math. Debrecen.* **6,** 290–297, (1959).
2. Palásti, I., On the Connectedness of Bichromatic Random Graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **8,** 431–441 (1963).
3. Roy, B., Cheminement et connexité dans l'application aux problèmes d'ordonancement. *Metra* **1,** (1962).
4. Palásti, I., On the Strong Connectedness of Directed Random Graphs. *Studia Scientiarum Mathematicarum Hungarica* **1–2,** 205–214 (1966).

DEFINITION OF MOST FAVOURABLE CHARACTERISTICS OF TRICOT FABRICS PRODUCED ON INTERLOCK MACHINES

by

N. RANCU, L. TÖVISSI and I. TEODORESCU

Academy of Economic Studies
Bucharest, Rumania

The following problems have emerged in the course of production.

(a) How to stabilize knitting parameters of tricot fabrics differing from one another in fineness and quality of thread (carded, combed) in order to reduce lasting deformation as far as possible during the processing and to reduce divergences in measures after washing, to a minimum.

(b) How to define technological procedure and final dressing in such a way as to achieve the highest possible dimensional stability of the final product (tricot) after washing and wearing.

To this end we examined 23 cotton tricot articles produced on interlock machines, differing from one another in fineness and quality of thread as well as to contexture.

During the first step, in the micro-test phase, we ascertained the relative knitting parameters for every single article. Knitting was performed in several variants for every single article and the best variant was chosen. In the second phase the final dressing parameters of tricot produced according to the best knitting procedure were fixed (semi-industrial knitting phase).

Taking the basic parameter characteristic for tricot of a given thread and contexture as the length of thread within one mesh or practically vertical density, we proposed to determine the best vertical density during the first phase. We therefore had every single article knitted in five or three steps of vertical density.

The tricot fabrics were analyzed in the laboratory during the following stages:

(a) Grey, in relaxed condition, before washing

(b) Grey, in relaxed condition, after washing

(c) Dyed, before washing, in case of tricot articles dyed individually before washing

(d) Dyed, after washing

The following characteristics were determined:

1. Longitudinal density (Dv) and transversal density (Do) (20 by 20 measurements),

2. Lasting horizontal stretch measured after the dynamometer test at a load of 1000 gf and 2000 gf and after a rest of 24 hours (10 by 10 measurement).

3. Lasting longitudinal stretch measured after the dynamometer test at a load of 3000 gf and a rest of 24 hours (10 by 10 measurements).

4. Dimensional charges in both directions after repeated washing of tricot fabrics (9–15 measurements per article).

5. Tne best Dv value variant was selected in the light of the laboratory examination results.

During the second phase certain preliminary experiments served to define how the technological procedure can be continued to the stage of calendering, in order to secure the least possible stretch on the tricot and to reduce its deformation to a minimum.

Tnis has been achieved by observing changes in the following parameters:

(a) Longitudinal and transversal density.

(b) Longitudinal variations of measure, fixed by way of marks of 500 mm previously applied on the raw, relaxed tricot at distances of 3 metres.

(c) Changes in the width of the tricot fabric.

Tne measurings were performed after the following procedures: Centrifuging, wet calendering and drying, final calendering and relaxing.

1. In order to find out the best longitudinal density of every single article we applied the following methods:

(a) Examination of the correlation between density and lasting stretch, performed after the mechanical processing and repeated washing of the tricot fabric.

(b) *t*-test.

The first method will be checked on an article produced from 18·5 tex(Nm 54) combed cotton yarn and of simple interlock contexture.

We set up a correlation chart of longitudinal density and of the following variants for all the eleven possible correlations:

(1) Longitudinal variation of measure (stretch) of grey, relaxed tricot after washing.

(2) Transversal variation of measure (stretch) of grey, relaxed tricot after washing.

(3) Lasting vertical variation of measure (stretch) of grey, relaxed tricot at a load of 3000 gf, before washing.

(4) Lasting longitudinal variation of measure (stretch) of grey, relaxed tricot at a load of 3000 gf, after washing.

(5) Lasting longitudinal variation of measure (stretch) of dyed tricot at a load of 3000 gf, before washing.

(6) Lasting transversal variation of measure (stretch) of grey, relaxed tricot at a load of 1000 gf, before washing.

(7) Lasting transversal variation of measure (stretch) of grey, relaxed tricot at a load of 1000 gf, after washing.

(8) Lasting transversal variation of measure (stretch) of dyed tricot at a load of 1000 gf, before washing.

(9) Lasting transversal variation of measure (stretch) of grey, relaxed tricot at a load of 2000 gf, before washing.

(10) Lasting transversal variation of measure (stretch) of grey, relaxed tricot at a load of 2000 gf, after washing.

(11) Lasting transversal variation of measure (stretch) of dyed tricot at a load of 2000 gf, before washing.

For every one of the above correlations a scattering diagram has been designed in order to illustrate the nature of the connection between the different characteristics.

An analysis of the scattering diagrams has permitted the inference that in the case of correlations (1), (3), (4), (5), (6), (7), (9) and (10), the regression may be considered linear, while in the case of correlations (2), (8) and (11) the regression curve is logarithmic. In certain cases — (1), (3), (5), (6), (7), (9) and (10) — doubts emerged in respect of the shape of the curve, therefore linearity tests were performed which showed that the regression was linear.

The equations and the equation system were constructed with the help of the principle of the least squares and solved according to the scheme of V. S. Nemtchinov. For the relation between longitudinal density and lasting transversal variation of measure of grey, relaxed tricot at a load of 1000 gf, after washing, the following equations resulted:

$$Y = 161{\cdot}86 - 0{\cdot}56179x.$$

In order to ascertain the intensity of the connection we then calculated the coefficient of correlation

$$r_{xy} = -0{\cdot}708.$$

A strong negative correlation was found to exist between the two variables examined.

Having applied the same procedure to the other variables we obtained the following regression equations:

$$\begin{aligned}
y_1 &= -\ 37{\cdot}7464 + \ 0{\cdot}4097x,\\
y_2 &= \ 285{\cdot}782 - 152{\cdot}631x,\\
y_3 &= \ 97{\cdot}235 + \ 0{\cdot}30265x,\\
y_4 &- \text{no correlation}\\
y_5 &= \ 114{\cdot}08 - \ 0{\cdot}022522x,\\
y_6 &= \ 190{\cdot}203 - \ 0{\cdot}9876x,\\
y_7 &= \ 161{\cdot}86 - \ 0{\cdot}56179x,\\
y_8 &= \ 4{\cdot}0796 - \ 1{\cdot}0774x,\\
y_9 &= \ 227{\cdot}285 - \ 1{\cdot}3145x,\\
y_{10} &= \ 183{\cdot}495 - \ 0{\cdot}72686x,\\
y_{11} &= \ 4{\cdot}3434 - \ 1{\cdot}181x.
\end{aligned}$$

The average longitudinal density which results in the smallest possible variation of measure after repeated washing and in the least lasting stretch after processing on the machine, as compared with the average values of the characteristics that have been examined, may be defined in the regression equation as ordinate value belonging to the abscissa.

It has been established that in all cases the density values of the different characteristics are very near to one another. All of them will be found between the values of 68·87 and 70·8 and around the mean value of $\bar{x}_{dr} = 69{\cdot}28$. We have thus found that for the articles under examination the vertical density of tricot produced on simple interlock machines from combed cotton yarn of 18·5 tex(Nm 54) fineness is 69 meshes per 5 cm.

The adequate vertical density of the other 22 articles under examination has been determined in a similar way.

2. In order to secure the necessary density and to enable us to work with the adequate average values, it is necessary to fix the tolerance limit which will have to be observed during production.

Tolerance limits for every single article were calculated according to the variation of value. To this end we defined for every one of the 23 articles under examination the variance for every density value examined per article. We arranged the variance squares according to the three contextures under examination: Tricot on simple interlock, tricot on interlock with fang effect and tricot on interlock with modified cam of the Overnit type.

In all groups the homogeneity of the variances were checked by way of the χ^2 test. In the case of simple interlock tricot we found that with the exception of one — which we eliminated — the variances were estimations of the same common σ^2-s variance. The value of the common variance was taken to be $s^2 = 1{\cdot}83$, which figure we obtained as the mean value of the remaining twelve variances.

The tolerance limits were fixed on a $P = 0{\cdot}95$ probability level.

These limits were defined with the help of the relation $q = 1 - 2\Phi(z)$, z being $\frac{\delta}{\sigma}$ and δ the semi-tolerance field, while σ is the standard deviation estimated from the experimental data.

The relation is valid if the distribution of knitting density is normal. The conditions under which this knitting density is being brought about, suggest the hypothesis of normal distribution.

The semi-tolerance field of meshes over a length of 5 cm is in the case of simple interlock tricot

$$\delta = 2{\cdot}64 \approx 2{\cdot}5,$$

in the case of interlock tricot with fang effect

$$\delta = 2{\cdot}82 \approx 2{\cdot}5$$

and in the case of interlock Overnit type tricot

$$\delta = 2{\cdot}68 \approx 2{\cdot}5.$$

As a result the tolerance field of all articles will be 5 meshes per 5 mm. In order to obtain tricot fabrics of good measure stability, the tolerance limits — as compared to the mean values — will not be fixed symmetrically, but 40 % will be applied to the lower and 60 % to the upper limit.

In the case of simple interlock tricot produced from 18·5 tex(Nm 54/1) combed cotton yarn, the tolerance limits of grey, relaxed tricot were

$$T_i = 69 - 2 = 67$$

and

$$T_s = 69 + 3 = 72$$

rows of meshes per 5 cm.

3. In order to secure tricot of such quality, the grey, relaxed tricot is controlled with the help of mathematical statistical methods. One bale of a lot of 25 is selected for five measurings. The mean value of the five measurings must remain within the limits of

$$\text{LCS} = M_0 + z_{q_I} \frac{\sigma}{\sqrt{n}}$$

and

$$\text{LCS} = M_0 - z_{q_I} \frac{\sigma}{\sqrt{n}},$$

where M_0 = the centre of the tolerance field
δ = the experimental standard deviation
n = the number of measurings
z_{q_I} = the risk of the product being judged defective, though it is not.

If q_I is 0·02, z will be 2·33, thus the tolerance limits for the article figuring in our example will be

$$\text{LCS} = 69{\cdot}5 + 2{\cdot}33 \frac{1{\cdot}35}{\sqrt{5}} = 71 \text{ meshes per 5 cm and}$$

$$\text{LCI} = 69{\cdot}5 - 2{\cdot}33 \frac{1{\cdot}35}{\sqrt{5}} = 68 \text{ meshes per 5 cm.}$$

4. The operative criterion for the stabilizing of calendering frame width in case of dry and wet calendering will be supplied by the data concerning the width of the tricot. The parameter of "tricot width" will also serve as a means of checking in respect of the procedure of final dressing in a wet state and in the drying of tricot.

Experiments were performed on machines of 18″, 22″ and 26″ diameter. The width of tricot fabrics produced at every machine width was measured in the following states: Grey and relaxed, dyed and centrifuged, wet calendered and dried, calendered and definitely dried. For every diameter six samples were measured and ten measurings were performed per sample.

The variances of the 18″ diameter did not significantly differ from one another. The tolerance limits in grey, relaxed state at a probability level of $P - 0{\cdot}95$ were

$$T_i = M_0 - z_{0{\cdot}05}\,\sigma = 472 \text{ mm},$$

$$T_s = M_0 - z_{0{\cdot}05}\,\sigma = 508 \text{ mm}.$$

In order to secure tricot fabric of stable measure and of width varying within the given limits, it is necessary that in case of grey, relaxed tricot and at a probability level of $q_I = 0{\cdot}05$ the mean value of the five measurings vary within the control limits of LCI = 480 mm and LCS = 500 mm.

The tolerance and control limits of tricot produced on machines of 18″, 22″ and 26″ diameter are shown in Table I.

TABLE I

Phase of technological procedure	Diameter in inches	T_i	T_s	LCI	LCS
Grey, relaxed	18	472	508	480	500
	22	593	627	601	619
	26	716	744	723	737
Dyed and centrifuged	18	352	382	—	—
	22	443	471	—	—
	26	532	562	—	—
Wet calendered and dried	18	381	403	—	—
	22	477	499	—	—
	26	573	597	—	—
Calendered and definitely dried	18	460	480	465	475
	22	583	597	587	593
	26	700	720	706	714

5. It will occur that — though longitudinal density and calendering width have been found adequate — the width and density of the final product will not comply with requirements.

This will happen when in the course of processing the tricot fabric will undergo inadmissible changes of dimension; possibly during wet dressing, drying and final calendering, if these procedures are inadequately performed. In order to avoid such inadmissible changes, it will be necessary to introduce preventive checking. To this end the changes of measure during the three procedures were studied in the following way:

Six pieces of every single article were knitted and subjected to final dressing. On these articles the marks which had been applied at distances of 500 mm when the fabric was grey and relaxed, were checked for possible variations. On every piece ten marks were applied. Tricot fabrics processed at three machines of different diameter were examined. With the help of values thus obtained the variance per diameter and piece for every single procedure could be calculated. The variances are shown in Table II. An examination of variance per piece, then of variance per diameter during the same procedure will show that they do not differ significantly; the conclusion that these variances are estimations of the same general variance is based on this fact.

The following variances have been established for the three technological procedures: $s_1^2 = 98{\cdot}18$; $s_2^2 = 76{\cdot}67$ and $s_3^2 = 57{\cdot}42$ mm.

TABLE II

Number of pieces	18″ Diameter			22″ Diameter			26″ Diameter		
	Technological operation			Technological operation			Technological operation		
	Dyed and centri-fuged	Wet calender-ized, dried	Calender-ized, finally dried	Dyed and centri-fuged	Wet calender-ized, dried	Calender-ized, finally dried	Dyed and centri-fuged	Wet calender-ized, dried	Calender-ized, finally dried
1	103·57	41	87·87	110·77	78·22	38·72	105·70	0·04	149·33
2	136·44	67·78	68·62	171·40	83·16	58·50	74·30	5·61	29·44
3	100	35·6	35·3	60·20	44·04	28·20	59	7·80	14·00
4	60·27	39·44	19·11	68·93	68·77	22·71	45·77	3·77	27·22
5	118·44	251·44	91·76	245·80	12·93	20·33	134·33	0·66	62·94
6	144·94	363·33	251·2	83·00	160·13	33·48	62·26	10·23	110·88
Mean dispersion	110·61	116·6	72·97	123·35	74·71	33·65	60·6	5·35	65·63

With the help of these variances it is possible to define the stretch variations admissible at 0·05 probability level and measured by variation expressed in percentage of the marks applied at the beginning over a length of 500 mm.

Operation of dyeing and centrifuging:

$$A_s = z_q\, s = 1{\cdot}96 \times 9{\cdot}9 \approx 20 \text{ mm.}$$

Operation of calendering and final drying:

$$A_s = 1{\cdot}96 \times 7{\cdot}56 \approx 15 \text{ mm.}$$

Operation of wet calendering and drying:

$$A_s = 1{\cdot}96 \times 8{\cdot}75 \approx 17 \text{ mm.}$$

In order to keep variations within the admissible tolerance limits, the technological operations will be regulated with the help of control charts the limits of which will be

LC = 10 mm for the dying and centrifuging operation
LC = 9 mm for wet calendering and drying and
LC = 5·5 mm for calendering and final drying.

6. Calender frame width was fixed for three machine diameters and then 2–3 calender widths were figured out by experiment for every machine diameter (3 pieces of tricot for every calender width). In the choice of calender width the measures of grey, relaxed tricot as well as the variations of measure during the last wet dressing operation were taken into consideration.

Measure variations for the fixation of the best calendar frame width which will occur in finished tricot fabrics after washing under household conditions were taken into account. The washing test was performed with the help of a 600 mm × 600 mm large test piece on which a 500 mm × 500 mm large control square had been marked. Measure variations after washing repeated three times under household conditions were examined.

Experiments were performed with three tricot fabrics for every calender width and of each single tricot fabric three test pieces were chosen: on every test piece three longitudinal measurings and three transversal measurings were performed, thus 27 measurements data were obtained for every direction of tricot fabric.

In order to find out the best calender frame width, the correlation of this variable with the measure variations of the dressed and domestically washed tricot was examined.

E. g. the regression between calender width (x) and transversal measure variation after washing (y) is expressed by the equation

$$y = -1{\cdot}0105x + 0{\cdot}02069.$$

The best calender frame width has been fixed according to the mean value of the measure variations experienced in both directions on the tricot fabric.

It has thus been established that e.g. for tricot produced on 20'' machine the best calender width is between 639 mm and 643·5 mm. According to this procedure the best calender widths were figured out for all articles under examination. An analysis of the width of grey, relaxed tricot, of the width of the calender frame and of the width of dressed tricot permitted us to find out that there exists a linear connection between these characteristics and the diameter of the machine. The regression method has been used in order to characterize this connection.

The regression between the width of grey, relaxed tricot (l_r) and the diameter (D) is expressed by the equation

$$l_c = 29D - 26.$$

Similarly it has been established that the regression between the frame-width of the calender (l_r) and the diameter of the machine is

$$l_r = 32{\cdot}5D - 65,$$

and the regression between the width of the dressed tricot (l_f) and the diameter of the machine (D)

$$l_f = 29D - 45{\cdot}6.$$

The mean value of longitudinal density (D_v) and of transversal density (D_0) of the articles under examination could be defined in the light of the laboratory tests to which the dressed tricot was subjected.

For every single article the variance was defined in both directions. It has been established that the longitudinal density variances of all three tricot groups will not differ significantly within either group and that the difference between the groups is not significant either. On the contrary, the variances of transversal density were found to differ significantly from one another, when the three groups were compared.

As a result the permissible deviation from standards of longitudinal density is the same for all article at $q = 0{\cdot}05$ probability level.

$A_{D_l} = z_{0\cdot05}\,\sigma - 1\cdot96 \times 1\cdot74 = 3\cdot48 \approx 3\cdot5$ meshes per 5 cm. The tolerance field is $2\sigma = 2 \times 3\cdot5 = 7$ meshes per 5 cm; taking into consideration that low density exercises a bad effect on stability of measure, we take 60% of the tolerance field for upper deviations and 40% for lower deviations.

The permissible deviation from mean transversal tricot density at $q = 0\cdot05$ probability level is;

(a) in the case of the simple interlock tricot group
$A_{D_0} = z_{0\cdot05}\,\sigma = 1\cdot96 \times 1\cdot65 = 3\cdot4 \approx 3$ meshes per 5 cm,

(b) in the case of the fang-effect interlock tricot group
$A_{D_0} = z_{0\cdot05}\,\sigma = 1\cdot96 \times 1\cdot65 = 3\cdot4 \approx 3$ meshes per 5 cm and

(c) in the case of the Overnit type tricot group
$A_{D_0} = z_{0\cdot05}\sigma = 1\cdot96 \times 1\cdot32 = 2\cdot6 \approx 2\cdot5$ meshes per 5 cm.

In order to control transversal density the tolerance field is divided in the same way as in the previous case and to control the procedure, statistical control cards are being used, the control limits of which will be calculated in the usual way.

EIN PROBLEM DER LINEAREN REGULIERUNG BEI EINER NICHTSTATIONÄREN STÖRUNG VON POLYNOMFORM

von

Z. REŽNÝ

Staatliches Forschungsinstitut für Wärmetechnik
Prag, Tschechoslowakei

1. EINLEITUNG

In diesem Referat handelt es sich um das Problem eines linearen Regelkreises, der in Abb. 1 wiedergegeben ist. Das Element mit der Übertragungsfunktion 1 stellt eine idealisierte Fertigungseinrichtung dar. Die Folge $\eta = \{\eta_n\}$ der Werte des kontrollierten Merkmals auf den einzelnen Produkten stellt die regulierte Größe dar und ist der Abb. 1 gemäß durch die Differenz der Störung ξ und der Stellgröße χ gegeben. In der Rückkopplung ist ein diskretes lineares Glied — ein Regler — eingeschaltet, dessen Übertragungsfunktion durch $R(z)$ bezeichnet ist. Um den Regelkreis realisierbar zu machen, muß man in der Rückkopplungslinie eine Verzögerung von mindestens einem Schritt voraussetzen. Zu dieser minimalen Verzögerung tritt noch allgemein eine weitere Verzögerung von m Schritten ($m \geqq 0$) hinzu, die durch die Beförderung des Erzeugnisses zum Meßgerät, Messung usw. verursacht wird.

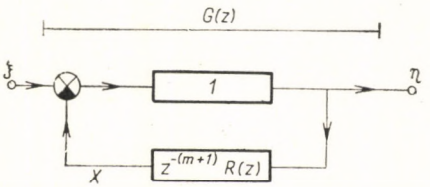


ABB. 1

L. Prouza hat in seiner Arbeit [1] die von A. Špaček ausgesprochene Hypothese bewiesen, daß das Problem der linearen Regulierung hier mit dem Problem der linearen Vorhersage (der Störgröße ξ durch die Größe χ) zusammenfällt. In der erwähnten Arbeit löste er die Aufgabe der Regulierung von minimalem mittlerem Quadrat $\mathbf{E}[(\eta_n - c)^2]$ mit der Nebenbedingung $\mathbf{E}(\eta_n) = c$ (c: geforderter konstanter Wert der Ausgangsgröße η), und zwar bei Abwesenheit der eigentlichen Verzögerung ($m = 0$). Die Störgröße ξ wurde von ihm als Summe einer Sprungfunktion und des weißen Geräusches vorausgesetzt; in einer späteren Arbeit [2] löste er dieselbe Aufgabe für die Summe einer linearen Funktion und des weißen Geräusches.

In dem vorliegenden Referat wird eine Lösung des Problems vorgeführt, die in den folgenden Richtungen verallgemeinert ist:

1. Es wird eine allgemeine eigentliche Verzögerung $m \geqq 0$ angenommen.
2. Die nichtstationäre Komponente der Störgröße wird in der Form eines Polynoms bekannten Grades mit zufälligen Koeffizienten angenommen.
3. Parallel mit dem Kriterium des minimalen mittleren Fehlerquadrates bei der Nebenbedingung $\mathbf{E}(\eta_n) = c$ betrachtet man dasselbe Kriterium ohne Nebenbedingung.

2. FORMULIERUNG DER AUSGANGSBEGRIFFE UND VORAUSSETZUNGEN

Die Eingangsstörgröße ξ des Kreises in Abb. 1 wird als eine Folge mit dem allgemeinen Element

$$\xi_n = \sum_{l=0}^{k} \alpha_l n^l + \gamma_n \tag{1}$$

angenommen; eine derartige Folge wird als *Eingang k-ten Grades* bezeichnet. Im Ausdruck (1) ist der Grad $k \geq 0$ des Störpolynoms im voraus bekannt und α_l, γ_n sind Zufallsveränderliche mit den Eigenschaften

$$0 < \mathbf{E}(\alpha_k^2) < +\infty, \tag{2}$$

$$\mathbf{E}(\gamma_n) = 0\,, \quad \mathbf{E}(\gamma_n \gamma_{n'}) = \sigma^2 \delta_{nn'}\,, \tag{3}$$

wobei σ^2 eine positive Konstante und δ das Kroneckersche Symbol bedeuten, und

$$\mathbf{E}(\alpha_l \gamma_n) = 0 \qquad (0 \leq l \leq k)\,. \tag{4}$$

Man kann den gemeinsamen Sollwert c aller Elemente der Ausgangsgröße η ohne Beschränkung der Allgemeinheit gleich Null setzen (dadurch wird nur eine Veränderung des Anfangspunktes des Maßstabes bei der Zufallsveränderlichen α_0 hervorgerufen).

Ähnlich wie es in [1] und [2] der Fall war, wird man von dem Regler in der Rückkopplung verlangen, daß dieser eine bestimmte endliche Einschwingzeit des Kreises sichere; diese praktische Forderung ist zugleich eine theoretische Notwendigkeit für die Lösbarkeit des Problems, was aus dem Folgenden ersichtlich wird. Aus dieser Forderung und aus den Operatorengleichungen des Kreises, die man nach der Abb. 1 aufstellen kann, wird man leicht feststellen, daß die Gesamtübertragungsfunktion des Kreises die Form

$$G(z) = 1 + \sum_{i=m+1}^{N} g_i z^{-i} \qquad (g_{m+1} \neq 0, g_N \neq 0) \tag{5}$$

haben soll, wo N der im voraus festgelegten, die Einschwingzeit angebenden Anzahl der Schritte gleicht. Die Übertragungsfunktion des Reglers wird dann nach Abb. 1

$$R(z) = z^{m+1}\left[\frac{1}{G(z)} - 1\right] \tag{6}$$

und die Übertragungsfunktion der Vorhersage (diejenige zwischen ξ als Eingangs- und χ als Ausgangsgröße) ist durch

$$A(z) = 1 - G(z) \tag{7}$$

gegeben. Daraus ist es ersichtlich, daß der Kreis durch seine Übertragungsfunktion $G(z)$, d. h. durch Elemente

$$g_i \quad (m+1 \leq i \leq N) \tag{8}$$

mathematisch beschrieben ist. Infolgedessen wird man im weiteren unter dem Begriff des »Kreises« eine Menge der Elemente (8) verstehen.

Die Anzahl der Elemente (8), $N - m$, wird als *Ordnung* des Kreises, m als seine *Verzögerung* und N als sein *Gedächtnis* bezeichnet. Die Ordnung des Kreises ist eine ganze nichtnegative Zahl; im Grenzfall $N = m$ hat man es mit einem trivialen Kreis von Nullordnung zu tun, dessen Übertragungsfunktion $G(z) = 1$ ist; daraus folgt $A(z) = R(z) = 0$; oder anders gesagt: der Kreis von Nullordnung wird am Ausgang die Eingangsstörung ohne Rückkopplungskorrektion (Abb. 2) reproduzieren.

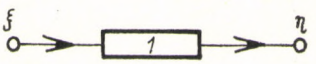


ABB. 2

3. k-STABILE UND OPTIMALE KREISE

Man nennt einen Kreis *k-stabil*, wenn die Folge $\{\mathbf{E}(\eta_n^2)\}_n$ der zweiten Momente von Elementen seiner Antwort auf den Eingang k-ten Grades (1) begrenzt ist. Es gilt dann der folgende

SATZ 3.1. *Notwendige und hinreichende Bedingungen der k-Stabilität sind*

$$\delta_{0j} + \sum_i{}' i^j g_i = 0 \quad \text{für} \quad 0 \leqq j \leqq k-1\,. \tag{9}$$

Sind diese Bedingungen erfüllt, dann gilt für die Antwort des Kreises auf den Eingang k-ten Grades (1)

$$\eta_n = (-1)^k (\delta_{0k} + \sum_i i^k g_i)\,\alpha_k + \gamma_n + \sum_i g_i \gamma_{n-i}\,, \tag{10}$$

hieraus

$$\mathbf{E}(\eta_n) = (-1)^k (\delta_{0k} + \sum_i i^k g_i)\,\mathbf{E}(\alpha_k)\,, \tag{11}$$

$$\mathbf{E}(\eta_n^2) = (\delta_{0k} + \sum_i i^k g_i)^2\,\mathbf{E}(\alpha_k^2) + (1 + \sum_i g_i^2)\,\sigma^2. \tag{12}$$

Aus diesem Satz (dessen Beweis, gleich wie die Beweise aller folgenden Sätze, im Interesse der Kontinuität des Textes, in den Nachtrag 4. verschoben wird) ergibt sich unter anderem, daß

(a) jeder Kreis 0-stabil ist;

(b) ein k-stabiler Kreis zugleich für alle $k' \leq k$ k'-stabil ist.

Die durch die Ordnung des Kreises gegebene Einschränkung ist im folgenden Satz enthalten:

SATZ 3.2. *Von den Kreisen der k-ten Ordnung mit der Verzögerung m ist eben einer k-stabil und keiner* $(k+1)$*-stabil.*

Nun werden wir uns mit den im Sinne dessen, was bereits in 1. angedeutet wurde, optimalen Kreisen befassen. Man definiere zu diesem Zweck zuerst:

(a) $\mathscr{M}(m, N)$ = die Menge der Kreise mit Verzögerung m und Gedächtnis N (für $m < N$) bzw. die Menge, die den Kreis von Nullordnung als ihr einziges Element enthält (für $m = N$);

(b) $\mathscr{M}_0(m, N, k)$ = die Menge der Kreise, die zu $\mathscr{M}(m, N)$ gehören und zugleich die Eigenschaft besitzen, daß für die Antwort η auf einen *beliebigen* Eingang k-ten Grades $\mathbf{E}(\eta_n) = 0$ gültig ist;

(c) $$\mathscr{M}^*(m, N) = \bigcup_{\alpha \geqq 0, \beta \geqq 0} \mathscr{M}(m + \alpha, N - \beta),$$

(d) $$\mathscr{M}_0^*(m, N, k) = \bigcup_{\alpha \geqq 0, \beta \geqq 0} \mathscr{M}_0(m + \alpha, N - \beta, k).$$

Zu jedem Paar (m, N) zieht man zweierlei *Optimalkreise* mit Bezug auf den Eingang k-ten Grades im Betracht: *Der Kreis vom Typus* Ik wird als ein Kreis aus der Menge $\mathscr{M}^*(m, N)$ mit minimalem mittlerem Quadrat $\mathbf{E}(\eta_n^2)$ der Antwort auf den Eingang k-ten Grades definiert. *Den Kreis vom Typus* IIk definiert man durch dieselbe Eigenschaft, aber in der Menge $\mathscr{M}_0^*(m,N,k)$. Die nachstehenden Sätze, die eine Bestimmung dieser Optimalkreise und ihre gegenseitigen Zusammenhänge enthalten, sind die eigentlichen Ergebnisse unserer Untersuchung.

SATZ 3.3 *Angenommen, daß* $N - m \geqq k$, *ist der Kreis vom Typus* Ik *für die gegebenen* m, N, k *durch die folgenden Bedingungen eindeutig bestimmt*:

1. *Er ist* k*-stabil.*
2. *Es existieren Konstanten* $d_0, d_1, \ldots, d_k$ *derart, daß für seine Elemente*

$$g_i = \sum_{l=0}^{k} d_l i^l \qquad (m + 1 \leq i \leq N) \tag{13}$$

und 3.

$$\delta_{0k} + \sum_i i^k g_i = -\varrho_k d_k \tag{14}$$

gilt, wobei

$$\varrho_k = \frac{\sigma^2}{\mathbf{E}(\alpha_k^2)}. \tag{15}$$

Durch die Bedingungen 1.–3. sind die Koeffizienten $d_0, \ldots, d_k$ *eindeutig bestimmt.*

SATZ 3.4. *Angenommen, daß* $N - m > k$, *ist der Kreis vom Typus* IIk *für die gegebenen* m, N, k *durch die folgenden Bedingungen eindeutig bestimmt*:

1. *Er ist* $(k + 1)$*-stabil.*
2. *Es existieren Konstanten* $d_0, d_1, \ldots, d_k$ *derart, daß* (13) *für seine Elemente gültig ist.*

Durch die Bedingungen 1., 2. sind die Koeffizienten $d_0, \ldots, d_k$ *eindeutig bestimmt.*

(Die Einschränkungen, durch die m, N, k gebunden sind, stellen Existenzbedingungen der entsprechenden Kreise dar, wie es aus dem Vergleich des Punktes 1. des Satzes 3.3 bzw. 3.4 mit dem Satz 3.2 zu sehen ist.)

Würden die Sätze 3.3, 3.4 als *Definitionen* der Optimalkreise angesehen, so würde der Kreis von Nullordnung als der Kreis vom Typus II(−1) für

beliebige m, N erscheinen. Für den derart erweiterten Komplex von Optimalkreisen gilt der nachstehende Satz, der einen interessanten Zusammenhang zwischen den Kreisen der beiden Typen angibt:

SATZ 3.5. *Ist $N - m > k$ und strebt ϱ_k gegen Null, dann konvergieren alle Elemente des Kreises vom Typus* Ik *gegen die entsprechenden Elemente des Kreises vom Typus* IIk.

Ist $N - m \geqq k$ und strebt ϱ_k gegen $+\infty$, dann konvergieren die Elemente des Kreises vom Typus Ik *gegen die entsprechenden Elemente des Kreises vom Typus* II$(k-1)$.

Werden die Sätze über die Optimalkreise zusammengefaßt, gelangt man zu folgender Regel für die Berechnung ihrer Elemente:

Man betrachte die Matrizen

$$\begin{aligned} A(m,N,k) &= \| (m+i)^{j-1} \| && (i = 1, \ldots, N-m; \\ & && j = 1, \ldots, k+1), \\ R(r,k) &= r. \| \delta_{i+j-2,2k} \| && (i,j = 1, \ldots, k+1), \\ S(m,N,k) &= \| s_{i+j-2}(m,N) \| && (i,j = 1, \ldots, k+1), \\ v(k) &= \| -1, 0, \ldots, 0 \|^{\mathrm{T}} && (\text{mit } k \text{ Nullelementen}), \end{aligned} \tag{16}$$

wobei i bzw. j die Zeilen bzw. die Spalten der Matrix bezeichnet, r eine reelle Zahl ist und

$$s_l(m,N) = \sum_{n=m+1}^{N} n^l \qquad (l \geqq 0), \tag{17}$$

welche Definition mit der leicht beweisbaren Rekursionsformel

$$s_l(m,N) = \frac{1}{l+1}\Big[(N+1)^{l+1} - (m+1)^{l+1} - \sum_{\alpha=0}^{l-1} \binom{l+1}{\alpha} s_\alpha(m,N)\Big] \tag{18}$$
$$(l \geqq 0)$$

äquivalent ist. Es sei endlich

$$g(m,N,k,r) = A(m,N,k)\big[S(m,N,k) + R(r,k)\big]^{-1} v(k). \tag{19}$$

Dann ist der Vektor $\| g_{m+1}, \ldots, g_N \|^{\mathrm{T}}$ der Elemente des Optimalkreises

$$g(m,N,k,\varrho_k) \qquad \text{für den Kreis vom Typus I}k$$
$$(0 \leqq k \leqq N-m) \tag{20}$$

und

$$g(m,N,k,0) = \lim_{r\to\infty} g(m,N,k+1,r) \qquad \text{für den Kreis vom Typus II}k \tag{21}$$
$$(-1 \leqq k < N-m).$$

Für $k = -1$ hat freilich in (21) nur die rechte Seite Sinn.

Numeriert man die Elemente von g von $m+1$ bis N übereinstimmend mit den Indizes der Elemente des Kreises, so ergibt sich durch Einsetzen in (19) und gemäss (16) z. B. für $k = 0$ und 1

$$g_i(m, N, 0, r) = -\frac{1}{N - m + r}, \tag{22}$$

bzw.

$$g_i(m, N, 1, r) = \tag{23}$$

$$= \frac{\left[-\frac{1}{3}(N - m + 1)(2N + m + 1) + 2\frac{r}{N - m}\right] + (N + m + 1)(i - m)}{\binom{N - m + 1}{3} + 2r}$$

$$(m + 1 \leq i \leq N).$$

Mit diesen Ausdrücken kann man die Gültigkeit des Satzes 3.5 für $k = 0$ und 1 unmittelbar nachprüfen. Außerdem kann man sich durch (23) von der interessanten Tatsache überzeugen, daß z. B. ein Kreis vom Typus I 1 mit $m = 0$ für $N = 4$ derselbe wie für $N = 5$ ist, falls $\varrho_1 = 20$ ist. Es ist also nicht richtig, daß ein Kreis vom Typus Ik immer zu der Menge $\mathcal{M}(m, N)$ gehört, wie es intuitiv vielleicht zu erwarten wäre. Vom theoretischen Standpunkte aus gesehen wäre es interessant, die Frage zu untersuchen, ob ähnliche Fälle eintreten können, wo eine verringerte *Verzögerung* m keine Verbesserung des optimalen Kreises zur Folge hat. Für $k = 0$ und 1 kann solch ein Fall nicht eintreten. Es wäre deshalb erforderlich, entweder diesen bei irgend einem $k > 1$ herauszufinden, oder andererseits die Behauptung zu beweisen, daß er bei keinem k vorkommt.

Zur Konstruktion eines Kreises vom Typus Ik ist nach (20) die Kenntnis des durch die Beziehung (15) definierten Parameters ϱ_k erforderlich. Man kann ein adaptives Modell mit einer sukzessiv gebildeten statistischen Schätzung des Parameters ϱ_k entwerfen, dem sich die Übertragungsfunktion des Reglers anpassen würde (vergl. die Arbeiten von M. Ullrich [3], [4]). Im Rahmen der linearen Theorie mit konstanten (d. h. zeitunabhängigen) Elementen des Kreises sind wir aber auf eine Approximation des optimalen Kreises — z. B. durch den Vektor (19) mit irgendwie erwähltem r — verwiesen; im speziellen ist aus (21) zu sehen, daß die Kreise vom Typus IIk und II$(k - 1)$ auch Approximationen des Kreises vom Typus Ik sind (bei der Wahl $r = 0$ bzw. $r = \infty$). Der Wert $r = r'$ ist besser als $r = r''$, was man durch

$$r' \succ r'' \quad (m, N, k)$$

bezeichnen wird, falls der Ausdruck (12), in den für g_i die Elemente des Vektors $g(m, N, k, r)$ eingesetzt wurden, bei $r = r'$ einen kleineren Wert als bei $r = r''$ annimmt; sind die beiden Werte des Ausdrucks (12) einander

gleich, wird man schreiben: $r' \approx r''\ (m, N, k)$. Dann kann man sich von der Richtigkeit z. B. der nachstehenden Behauptungen überzeugen:

$$\text{Ist } r_{1,2} \neq -(N-m), \quad \text{dann hat man}$$

$$(N-m+\varrho_0)\left(N-m+\frac{r_1+r_2}{2}\right) < \text{bzw.} = (N-m+r_1)(N-m+r_2) \Leftrightarrow$$

$$\Leftrightarrow \min(r_1, r_2) \succ \text{bzw.} \approx \max(r_1, r_2) \quad (m, N, 0)\,.$$

Im speziellen:

(a) $$0 < r < +\infty \Rightarrow r \succ -r \quad (m, N, 0)\,;$$

(b) $$\varrho_0 \geqq N-m\,,\ 0 < r < +\infty \Rightarrow r \succ 0 \quad (m, N, 0)\,;$$

(c) $$\varrho_0 < N-m,\ 0 < r_1 < r_2 = \frac{2(N-m)\varrho_0}{N-m-\varrho_0} < r_3 \Rightarrow$$

$$\Rightarrow r_1 \succ 0 \approx r_2 \succ r_3 \quad (m, N, 0)\,;$$

(d) $$\varrho_0 \leqq N-m,\ 0 < r < +\infty \Rightarrow r \succ \infty \quad (m, N, 0)\,;$$

(e) $$\varrho_0 > N-m,\ 0 < r_1 < r_2 = \frac{\varrho_0 - (N-m)}{2} < r_3 < +\infty \Rightarrow$$

$$\Rightarrow r_1 \prec \infty \approx r_2 \prec r_3 \quad (m, N, 0)\,;$$

(f) $$\varrho_0 < \text{bzw.} = N-m \Leftrightarrow 0 \succ \text{bzw.} \approx \infty \quad (m, N, 0)\,.$$

Für den Fall $k = 1$ sei hier mindestens ein spezielles Ergebnis angeführt, welches die in [2] betrachteten Kreise betrifft:

$$\varrho_1 < \text{bzw.} = \frac{1}{2}\binom{N+1}{3} \Leftrightarrow 0 \succ \text{bzw.} \approx \infty \quad (0, N, 1)\,.$$

4. NACHTRAG-BEWEIS DER SÄTZE DES § 3.

Beweis des Satzes 3.1. Mit Bezug auf (5) gilt es für die Antwort η auf die Eingangsfolge ξ:

$$\eta_n = \xi_n + \sum_{i=m+1}^{N} g_i \xi_{n-i}\,. \tag{24}$$

Durch Einsetzen für ξ_n aus (1) und unter Anwendung von (2), (3) und (4) kann man sich leicht überzeugen, daß im Falle einer Nichterfüllung der Bedingungen (9) $\lim_{n\to\infty} \mathbf{E}(\eta_n^2) = +\infty$ ist, so daß der Kreis nicht k-stabil ist, und umgekehrt, daß aus den Bedingungen (9) sich (10) ergibt. Die Beziehungen für die Momente (11), (12) sind dann unmittelbare Folgen von (10).

BEWEIS DES SATZES 3.2. 1. Ist $N > m$, dann stellen die Bedingungen (9) ein System von k inhomogenen Gleichungen für $N - m$ Unbekannten mit der Matrix des Systems $A(m, N, k)$ dar (siehe (16)). Daraus ergibt sich leicht die Gültigkeit des Satzes. 2. Der Kreis von Nullordnung erfüllt nicht die Bedingungen (9) für jedes $k > 0$.

BEWEIS DES SATZES 3.3. Ist ein Kreis k-stabil, dann gilt (12), indem in umgekehrtem Falle $\mathbf{E}(\eta_n^2) \to +\infty$. Man muß also einen Kreis vom Typus I^k unter den k-stabilen Kreisen suchen, was eben die Bedingung 1. ist. Daraus ergibt sich nach dem Satz 3.2, daß im Falle $N - m < k$ kein Kreis vom Typus I^k vorhanden ist.

Es sind also die Eigenschaften 2., 3. unter der Voraussetzung $N - m \geqq k$ zu beweisen. Nach der Definition des Kreises vom Typus I^k und mit Bezug auf die Bedingung 1. minimalisiert der Kreis den Ausdruck (12) unter den Nebenbedingungen (9), d. h.

$$(\delta_{0k} + \sum_i i^k g_i)^2 \mathbf{E}(\alpha_k^2) + (1 + \sum_i g_i^2)\,\sigma^2 + \sum_{l=0}^{k-1} L_l \sum_i i^l g_i = \min . \tag{25}$$

Schreibt man die Lagrangeschen Konstanten in der Form

$$L_l = -2\sigma^2 d_l \qquad (0 \leqq l \leqq k-1),$$

dann erscheint die Bedingung, daß die partiellen Ableitungen nach allen g_i gleich Null sind, äquivalent mit dem System der Beziehungen (13), (14). Hinsichtlich der Existenz und Eindeutigkeit der Koeffizienten d_l reicht es nun aus, die Gleichung (13) mit (20) und (19) zu vergleichen, was

$$\| d_0, \ldots, d_k \|^{\mathrm{T}} = [S(m, N, k) + R(r, k)]^{-1} v(k) \,|_{r=\varrho_k} \tag{26}$$

ergibt. Bei $k = 0$ hat die Matrix $S(m, N, k) + R(r, k)$ ein einziges Element $s_0 + r = N - m + r$, das nur für $r = -(N - m)$ $(\leqq 0)$, also nicht für $r = \varrho_0$ (> 0), gleich Null ist. Die rechte Seite in (26) hat also in jedem Falle Sinn. — Bei $k \geqq 1$ ist die erwähnte Matrix nur für

$$r = -\frac{|\,S(m, N, k)\,|}{|\,S(m, N, k-1)\,|} \tag{27}$$

singulär; dieser Ausdruck, wie es aus dem nachstehenden Hilfssatz 4.1 folgt, ist gleichfalls nicht positiv und infolgedessen von ϱ_k verschieden.

HILFSSATZ 4.1. *Die Matrix $S(m, N, k)$ ist für $N - m = k \geqq 1$ positiv semidefinit und für $N - m > k \geqq 0$ positiv definit.*

BEWEIS. Man betrachte die quadratische Form Q eines beliebigen Vektors $x = \|x_0, \ldots, x_k\|^{\mathrm{T}}$ mit der Matrix $S(m, N, k)$. Es gilt dann

$$Q = x^{\mathrm{T}} S(m, N, k)\, x = \sum_{i=0}^{k} \sum_{j=0}^{k} s_{i+j}(m, N)\, x_i x_j =$$

$$= \sum_{i=0}^{k} \sum_{j=0}^{k} \sum_{n=m+1}^{N} n^{i+j}\, x_i x_j = \sum_{n=m+1}^{N} (\sum_{i=0}^{k} n^i x_i)^2 \geqq 0 .$$

Es gilt also $Q = 0$ genau dann, wenn

$$A(m, N, k)\, x = o \tag{28}$$

ist. Die Zahl der Elemente des Vektors x ist $k + 1$, der Rang der Matrix $A(m, N, k)$ ist $\min(N - m, k + 1)$. Das System der Gleichungen (28) hat also eine von Null verschiedene Lösung genau dann, wenn $N - m < k + 1$, d. h. $k \geqq N - m$.

Beweis des Satzes 3.4. Ein Kreis vom Typus IIk muß aus demselben Grunde k-stabil sein wie ein Kreis vom Typus Ik. Bei beliebigem Eingang k-ten Grades soll außerdem $\mathbf{E}(\eta_n) = 0$ gelten, was nach (11) mit äquivalent ist.

$$\delta_{0k} + \sum_i i^k g_i = 0 \tag{29}$$

Man sieht also, daß die Bedingungen (9) für $0 \leqq j \leqq k$ erfüllt sind, so daß der Kreis dem Satze 3.1 gemäß $(k + 1)$-stabil ist. Daraus ergibt sich nach dem Satze 3.2, daß $N - m < k + 1$ nicht bestehen kann. Die Eigenschaft 2. (vorausgesetzt, dass $N - m > k$ ist) wird man ähnlich wie die Eigenschaften 2., 3. Kreises vom Typus Ik beweisen. Ein einziger Unterschied besteht darin, daß das erste Glied im Ausdruck (25) jetzt (mit Bezug auf (29)) gleich Null ist und daß man in dem letzten über l von 0 bis k anstatt bis $k - 1$ summiert.

Die Bedingungen 1., 2. des Satzes 3.4 werden äquivalent mit den Bedingungen 1.—3. des Satzes 3.3, wenn man darin $\varrho_k = 0$ anstatt der Gleichung (15) schreibt. Dadurch bekommt man (mit Bezug auf die Stetigkeit) die erste Behauptung des Satzes 3.5; zugleich wird (26) in

$$\| d_0, \ldots, d_k \|^{\mathrm{T}} = S^{-1}(m, N, k)\, v(k) \tag{30}$$

übergehen. Die Matrix $S(m, N, k)$ ist nach dem Hilfssatze 4.1 positiv definit. Dadurch ist die Existenz und Eindeutigkeit des Vektors (30) bewiesen.

Beweis des Satzes 3.5. Die erste Behauptung wurde bereits oben bewiesen. Hinsichtlich des Restes genügt es, sich zu überzeugen, daß nach (26)

$$\lim_{\varrho_k \to \infty} d_k = 0$$

ist. Durch Vergleich von Sätzen 3.3 und 3.4 ist es nämlich sofort ersichtlich, daß man einen Kreis vom Typus II$(k - 1)$ bekommt, wenn die Bedingung 3. im Satz 3.3 durch die Bedingung $d_k = 0$ ersetzt wird.

5. SCHLUßBEMERKUNGEN

Bei der Formulierung der gelösten Aufgabe ging man vor allem von den praktischen Voraussetzungen aus, die für eine breite Klasse von Produktionsoperationen, besonders im Maschinenbau. als typisch erscheinen. Dies bezieht sich auf die geläufig verwendete Zerlegung der Störgröße in eine

»systematische« und eine »zufällige« Komponente. Für die erste davon wurde eine allgemeine Form eines Polynoms vorausgesetzt, die Erscheinungen wie den »Sprung« durch die Einwirkung einer geringen Veränderung von Eigenschaften des bearbeiteten Werkstoffes, oder die durch den Verschleiss des Werkzeuges verursachte lineare, quadratische usw. Zeitveränderung der bearbeiteten Dimension u. ä. einschließt. Die zweite Komponente der Störgröße ist durch die Schwankung des regulierten Merkmals gegeben, die durch die Wirkung einer Reihe von geringen Einflüsse wie Maschinenschwingungen, Inhomogenität des bearbeiteten Werkstoffes u. ä. verursacht wird.

Was die Eigenschaften des Störpolynoms anbelangt, erscheinen die praktischen Situationen durch die Voraussetzung der *Zufallskoeffizienten* am besten ausgedrückt. Man begegnet zwar selten Fällen, wo diese Koeffizienten beliebige Werte annehmen können, demgegenüber kann aber eine vollkommene Kenntnis *a priori* von deren Verteilung praktisch nicht vorausgesetzt werden. Es ist deshalb von Vorteil, daß sich ein Optimalregler nur von dem Verhältnis ϱ_k zweier stochastischer Parameter abhängig zeigt. Dadurch ist auch der im Text angedeutete Weg zu einer Näherungslösung der Optimalisation erschlossen. Durch die Zufallskoeffizienten wird außerdem auch die Möglichkeit der Veränderungen von Parametern der Störung im Verlauf des regulierten Prozesses berücksichtigt.

Zu den betrachteten Kriterien der Optimalität sei noch bemerkt, daß die Einhaltung der durch einen *Toleranzbereich* gegebenen technologischen Vorschrift in der technischen Praxis für die Qualität des Produktionsvorgangs am wichtigsten erscheint. Darum ging man an der ersten Stelle von der Forderung einer minimalen »Herstellungstoleranz« aus, die durch die Definition des Kreises vom Typus Ik ausgedrückt ist. Demgegenüber ergibt sich der Kreis vom Typus IIk aus den abweichend und mehr detailliert spezifizierten Forderungen an die Qualität der Regulierung; interessant ist die Feststellung, zu der man gelangt ist, nämlich daß dieser Kreis als ein spezieller Fall der Abschätzung einer einfachen mittleren quadratischen Optimallösung, d. h. des Kreises vom Typus Ik bzw. I$(k + 1)$, erscheint.

Eine Realisierung der an dieser Stelle betrachteten Kreise mit Hilfe einer digitalen Rechenanlage ist nach dem allgemeinen Schema § 4.7 in [5], durchführbar (vgl. auch [1], Abb. 2).

LITERATUR

1. PROUZA, L., Zur linearen Theorie automatischer Einstellvorrichtungen. *Aplikace matematiky* **5**, Nr. **3**, 196–201 (1960).
2. PROUZA, L., Zur linearen Theorie automatischer Einstellvorrichtungen. *Abhandlungen der Konferenz über die Anwendungen der mathematischen Statistik im Maschinenbau*, 139–144, Prag (1962).
3. ULLRICH, M., Optimum Control of Some Stochastic Systems. *Proc. of the VIII ETAN Conference*, Agram (1963)
4. ULLRICH, M., Ein Wahrscheinlichkeitszutritt zur Bestimmung von Optimalreglern. *Jahrbuch »Problémy kybernetiky«*, 140—146, Prag (1965).
5. RAGAZZINI, J. R., FRANKLIN, G. F., Sampled-data Control Systems. McGraw-Hill, New York (1958).

INFORMATION AND STATISTICS

by

A. RÉNYI

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary

To anybody familiar with the principles of information theory it should be evident that statistical methods can be considered from the point of view of how they handle the available *amount* of information. It seems to the author that in current statistical thinking this point of view is not as consequently applied as it deserves. The reason for this seems to be that if one accepts as a leading principle in statistics to look at everything from the point of view of information theory, one is forced to accept the Bayesian point of view, and the amount of relevant information from sampling can be measured only if the prior information is taken into account, because the amount of information concerning the parameter obtained from the sample can be computed only by comparing the prior and posterior distributions of the parameter.

In what follows we try to outline some fundamental questions of the information theoretical approach to certain basic problems of statistics.

Before going into details we have to define certain notions of information theory. If ξ and ϑ are discrete random variables defined on the same probability space, ξ taking on the different values $a_1, a_2, \ldots$ and ϑ the values $b_1, b_2, \ldots$ *the amount of information contained in the value of* ξ *with respect to* ϑ is given by

$$\mathbf{I}(\xi, \vartheta) = \sum_{j,k} r_{jk} \log \frac{r_{jk}}{p_j q_k}, \tag{1}$$

where

$$r_{jk} = \mathbf{P}(\xi = a_j, \vartheta = b_k), \; p_j = \mathbf{P}(\xi = a_j), \; q_k = \mathbf{P}(\vartheta = b_k).$$

The *gain of information* concerning a discrete random variable ζ obtained from observing the event 0, if the prior distribution of ζ before taking the observations was $\mathbf{P} = \{p_1, p_2, \ldots\}$ where $p_i = \mathbf{P}(\zeta = Z_i)$ and the posterior distribution of ζ after observing 0 is $\mathbf{Q} = \{q_1, q_2, \ldots\}$ where $q_i = \mathbf{P}(\zeta = Z_i)$ is given by

$$\mathbf{I}\langle \mathbf{Q} \,||\, \mathbf{P} \rangle = \sum_i q_i \log \frac{q_i}{p_i}. \tag{2}$$

The entropy of a discrete distribution $\mathbf{P} = \{p_1, p_2, \ldots\}$ is defined by

$$\mathbf{H}(\mathbf{P}) = \sum_k p_k \log \frac{1}{p_k}. \tag{3}$$

Now let us consider the following basic problem of statistics: We observe a sample, i.e. a random vector ξ the distribution of which depends on the parameter ϑ (which may be a single number or a vector) then in case ϑ itself is a discrete valued random variable, the amount of information in the sample ξ concerning the parameter ϑ is given — in case ξ is also discrete valued — by $\mathbf{I}(\xi, \vartheta)$ as defined by (1). This quantity can be expressed also as

$$\mathbf{I}(\xi, \vartheta) = \mathbf{E}(\mathbf{I}\langle \mathbf{Q}_\xi \,||\, \mathbf{P}\rangle), \tag{4}$$

where $\mathbf{P}$ is the prior distribution of ϑ and $\mathbf{Q}_\xi$ its posterior distribution after observing the sample ξ, computed according to Bayes's theorem and $\mathbf{I} < \mathbf{Q}_\xi \,||\, \mathbf{P} >$ is defined by (2), $\mathbf{Q}_\xi$ depends of course on the sample and thus $\mathbf{I} < \mathbf{Q}_\xi \,||\, \mathbf{P} >$ is itself a random variable, and $\mathbf{I}(\xi, \vartheta)$ is equal to its expectation. Formula (4) may serve as the definition of $\mathbf{I}(\xi, \vartheta)$ in case ξ has an arbitrary (not discrete) distribution. A third expression of $\mathbf{I}(\xi, \vartheta)$ is

$$\mathbf{I}(\xi, \vartheta) = \mathbf{H}(\vartheta) - \mathbf{E}(\mathbf{H}(\vartheta \mid \xi)), \tag{5}$$

where $\mathbf{H}(\vartheta)$ and $\mathbf{H}(\vartheta \mid \xi)$ denote the entropy of the prior resp. posterior distribution of ϑ. Formula (5) is valid in the general case (i.e. not only for discrete-valued random variables). The use of the quantity $\mathbf{I}(\xi, \vartheta)$ in this context is quite natural and was suggested by several statisticians (see e.g. D. V. Lindley [1], where references to previous literature are also given), but it was not very frequently used up to now. The evaluation of the amount of information in a sample concerning the parameter is very useful e.g. in the design of experiments as it helps in choosing that alternative which gives the maximal amount of information, or to compare different sampling plans from the same point of view. If one chooses a statistic $S(\xi)$, the ratio

$$\frac{\mathbf{I}(S(\xi), \vartheta)}{\mathbf{I}(\xi, \vartheta)} \tag{6}$$

shows how large a part of the available information is really utilized by using only the statistic $S(\xi)$; it is easy to show that the ratio (6) is equal to 1 if and only if $S(\xi)$ is a *sufficient statistic for* ϑ. If an infinite sequence of observations $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ is taken and we use the notation $\xi^{(n)} = (\xi_1, \xi_2, \ldots, \xi_n)$ for the set of the first n observations, the sequence of random variables $I < \mathbf{Q}_{\xi^{(n)}}, \mathbf{P} >$ is a semi-martingale. If the random variables ξ_k are independent and identically distributed for each fixed value of the parameter (in which case the observations ξ_k are equivalent random variables) and these distributions are different for different values of the parameter, then if ϑ is discrete valued, one has

$$\lim_{n\to\infty} \mathbf{I}(\xi^{(n)}, \vartheta) = \mathbf{H}(\vartheta); \tag{7}$$

moreover the (non-negative) difference

$$\mathbf{H}(\vartheta) - \mathbf{I}(\xi^{(n)}, \vartheta) \tag{8}$$

tends to 0 for $n \to \infty$ exponentially fast (i.e. is of order $O(q^n)$ where $0 < q < 1$).

As an example for the use of the information-theoretical point of view let us consider the case when ϑ has only two possible values ϑ_0 and ϑ_1. Let us denote the hypothesis $\vartheta = \vartheta_i$ by $H_i (i = 0, 1)$. If in the sample space a critical region A is chosen and the statistic $\alpha = \alpha(\xi)$ is defined by $\alpha = 1$ if $\xi \in A$ and $\alpha = 0$ if $\xi \notin A$, then $\mathbf{I}(\alpha, \vartheta)$ gives us the amount of information contained in the observation (of the sample lying in the critical region or not), concerning the question whether H_0 or H_1 is true.

According to the above point of view we should choose such a region A as the critical region for which $\mathbf{I}(\alpha, \vartheta)$ is maximal; thus $\mathbf{I}(\alpha, \vartheta)$ acts like a loss-function. As it is well known, while in industrial applications the loss function can often be determined it usually cannot be done if the purpose of the statistical analysis is scientific. In such cases the measurement of the amount of information replaces the loss function, and eliminates the arbitrariness in the choice of the level of the error of the first kind. However, the best choice of the critical region depends of course on the prior probability $w_0 = \mathbf{P}(H_0)$ of the hypothesis H_0; to make the problem definite, one has to know the value of w_0. If w_0 is unknown, one has to choose for w_0 some value in order to get a definite test. Thus the arbitrariness which was thrown out concerning the choice of the error of the first kind returns in certain cases by another door, namely when the problem is such that w_0 is unknown. However, the dependence of the critical region on a slight change of w_0 is usually not strong and thus a reasonable estimate of w_0 is practically sufficient.

Finally it should be added that while most statisticians who advocate the Bayesian point of view consider the prior distribution as necessarily of a subjective nature, expressing the experimenters beliefs, in many cases a prior distribution having an objective meaning can — at least approximately — be given, e.g. by evaluation of some preliminary experiments as is shown by the results of H. Robbins concerning what he calls the "empirical Bayes approach". A highly illuminating discussion of this point of view was recently given by J. Neyman [2].

The proofs of the results announced here, and further developments of these ideas are given elsewhere (see: [3], [4], [5], [6], [7] [8]).

REFERENCES

1. Lindley, D. V., On a Measure of the Information Provided by an Experiment. *Annals of Math. Stat.* **27,** 986–1005 (1956).
2. Neyman, J., Two Breakthroughs in the Theory of Statistical Decision Making. *Revue de l'Inst. Internat. de Statistique* **30,** 11–27 (1962).
3. Rényi, A., On the Amount of Information Concerning an Unknown Parameter in a Sequence of Observations. *Publ. Math. Inst. Hung. Acad. Sci.* **9,** 617–624 (1964).
4. Rényi, A., On the Amount of Missing Information and the Neyman–Pearson Lemma. *Research Papers in Statistics, Festschrift for J. Neyman,* Wiley, London, 281—288 (1966).
5. Rényi, A., On some Basic Problems of Statistics from the Point of View of Information Theory. *Proceedings of the 5th Berkeley Symposium* (in print).
6. Rényi, A., On the Amount of Information in a Frequency Count. *35th Session of the ISI Beograd, 1965.*
7. Rényi, A., On the Amount of Information in a Random Variable Concerning an Event. *Journal of Mathematical Sciences (Delhi)* **1,** 30–33 (1966).
8. Rényi, A., Statistics and Information Theory. *Studia Sci. Math. Acad. Sci. Hung.* **2,** 249—256 (1967).

SOME REMARKS ON THE ESTIMATION OF FRACTION DEFECTIVE AND RELIABILITY IN THE CASE OF NORMAL DISTRIBUTION

by

K. SARKADI

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary

Let $x_1, \ldots, x_n$ be a random sample from a normal distribution with expectation μ and variance σ^2. For the quantity $p = \mathbf{P}(x_1 < L)$, where L is a constant, (μ unknown, σ known or unknown), mainly two types of estimators are in common use. Let us denote by $\bar{x}$ and s the sample mean and variance, i.e.:

$$\bar{x} = \frac{x_1 + \ldots + x_n}{n},$$

$$s^2 = \frac{(x_1 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n-1}.$$

The principle of equating moments leads to the estimator

$$t_1 = \Phi\left(\frac{L - \bar{x}}{\sigma}\right)$$

in the case of known variance, and to the estimator

$$t_2 = \Phi\left(\frac{L - \bar{x}}{s}\right)$$

in the case of unknown variance.

Here $\Phi(x)$ is the standardized normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

The estimators t_1 and t_2 are biased.

Lieberman and Resnikov (p. 465 in [3]) have given the best unbiased estimator of p in case of known variance:

$$t_3 = \mathbf{E}(x_1 - L \mid \bar{x}) = \Phi\left(\frac{L - \bar{x}}{\sigma}\sqrt{\frac{n}{n-1}}\right).$$

TABLE I

This table gives the minimum variance unbiased estimator of $t_4 = P(x_1 < L)$ or $P(x_1 > U)$ in function of $C = C_e$ or C_u where $C_e = \frac{x - L}{s}$, $C_u = \frac{U - x}{s}$

Entries in the table are given in per cent

C	n											
	3	4	5	6	7	8	9	10	11	12	13	14
0	50	50	50	50	50	50	50	50	50	50	50	50
0·1	47	47	46	46	46	46	46	46	46	46	46	46
0·2	45	43	43	43	43	42	42	42	42	42	42	42
0·3	42	40	39	39	39	39	39	39	39	38	38	38
0·4	39	37	36	35	35	35	35	35	35	35	35	35
0·5	36	33	32	32	32	32	31	31	31	31	31	31
0·6	33	30	29	28	28	28	28	28	28	28	28	28
0·7	29	26	26	25	25	25	25	25	25	25	25	24
0·8	26	23	23	22	22	22	22	22	22	21	21	21
0·9	22	20	19	19	19	19	19	19	19	19	19	19
1·0	17	17	16	16	16	16	16	16	16	16	16	16
1·1	10	13	13	13	13	14	14	14	14	14	14	14
1·2	0·00	10	11	11	11	11	11	11	11	11	11	11
1·3	0·00	6·7	8·2	8·7	8·9	9·1	9·1	9·2	9·2	9·3	9·3	9·4
1·4	0·00	3·3	5·9	6·7	7·0	7·2	7·3	7·4	7·5	7·5	7·6	7·6
1·5	0·00	0·00	3·8	4·8	5·3	5·6	5·7	5·9	6·0	6·1	6·1	6·2
1·6	0·00	0·00	2·0	3·2	3·8	4·2	4·4	4·5	4·6	4·7	4·8	4·9
1·7	0·00	0·00	0·66	2·0	2·6	3·0	3·2	3·4	3·5	3·6	3·7	3·8
1·8	0·00	0·00	0·60	1·0	1·7	2·1	2·3	2·5	2·6	2·7	2·8	2·9
1·9	0·00	0·00	0·00	0·36	0·93	1·3	1·6	1·8	1·9	2·0	2·1	2·2
2·0	0·00	0·00	0·00	0·02	0·43	0·77	1·0	1·2	1·3	1·4	1·5	1·6
2·1	0·00	0·00	0·00	0·00	0·14	0·39	0·59	0·74	0·86	0·96	1·0	1·1
2·2	0·00	0·00	0·00	0·00	0·02	0·16	0·32	0·44	0·54	0·63	0·70	0·76
2·3	0·00	0·00	0·00	0·00	0·00	0·04	0·14	0·23	0·32	0·39	0·45	0·50
2·4	0·00	0 ·00	0·00	0·00	0·00	0·00	0·05	0·11	0·17	0·22	0·27	0·31
2·5	0·00	0·00	0·00	0·00	0·00	0·00	0·01	0·04	0·08	0·12	0·16	0·19
2·6	0 00	0·00	0·00	0·00	0·00	0·00	0 00	0·00	0·03	0·06	0·08	0·11
2·7	0·00	0·00	0·00	0·00	0·00	0·00	0 00	0·00	0·01	0·02	0·04	0·05
2·8	0·00	0·00	0·00	0·00	0·00	0·00	0 00	0·00	0·00	0·01	0·02	0·02
2·9	0·00	0·00	0·00	0·00	0·00	0·00	0 00	0·00	0·00	0·00	0·01	0·01
3·0	0·00	0·00	0·00	0·00	0·00	0·00	0 00	0·00	0·00	0·00	0·00	0·00
3·1	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
3·2	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
3·3	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
3·4	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
3 5	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
3·6	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00
3·7	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00

$$t_4(-C) = 100\% - t_4(C)$$

TABLE I (continued)

This table gives the minimum variance unbiased estimator of $t_4 = P(x_1 < L)$ or $P(x_1 > U)$ in function of $C = C_e$ or C_u where $C_e = \frac{\bar{x} - L}{s}$, $C_u = \frac{u - x}{s}$

Entries in the table are given in per cent

C	n											
	15	16	17	18	19	20	25	30	35	40	50	∞
0	50	50	50	50	50	50	50	50	50	50	50	50
0·1	46	46	46	46	46	46	46	46	46	46	46	46
0·2	42	42	42	42	42	42	42	42	42	42	42	42
0·3	38	38	38	38	38	38	38	38	38	38	38	38
0·4	35	35	35	35	35	35	35	35	35	35	35	34
0·5	31	31	31	31	31	31	31	31	31	31	31	31
0·6	28	28	28	28	28	28	28	28	28	28	28	27
0·7	24	24	24	24	24	24	24	24	24	24	24	24
0·8	21	21	21	21	21	21	21	21	21	21	21	21
0·9	19	19	19	19	19	19	18	18	18	18	18	18
1·0	16	16	16	16	16	16	16	16	16	16	16	16
1·1	14	14	14	14	14	14	14	14	14	14	14	14
1·2	11	11	11	11	11	11	11	11	11	11	11	12
1·3	9·4	9·4	9·4	9·4	9·5	9·5	9·6	9·6	9·6	9·6	9·6	9·7
1·4	7·7	7·8	7·8	7·8	7·8	7·8	7·9	7·9	7·9	7·9	8·0	8·1
1·5	6·2	6·2	6·3	6·3	6·3	6·3	6·4	6·5	6·5	6·5	6·5	6·7
1·6	4·9	5·0	5·0	5·0	5·1	5·1	5·2	5·2	5·3	5·3	5·3	5·5
1·7	3·8	3·9	3·9	4·0	4·0	4·0	4·1	4·2	4·2	4·2	4·3	4·5
1·8	2·9	3·0	3·0	3·1	3·1	3·1	3·2	3·3	3·3	3·4	3·5	3·6
1·9	2·2	2·3	2·3	2·4	2·4	2·4	2·5	2·5	2·6	2·6	2·7	2·9
2·0	1·6	1·6	1·7	1·7	1·8	1·8	1·9	2·0	2·0	2·1	2·1	2·3
2·1	1·2	1·2	1·2	1·3	1·3	1·3	1·4	1·5	1·5	1·6	1·6	1·8
2·2	0·80	0·85	0·90	0·91	0·94	0·97	1·1	1·1	1·2	1·2	1·3	1·4
2·3	0·54	0·58	0·60	0·64	0·66	0·68	0·77	0·82	0·86	0·89	0·93	1·1
2·4	0·35	0·38	0·41	0·43	0·46	0·47	0·55	0·60	0·62	0·65	0·69	0·82
2·5	0·21	0·24	0·27	0·28	0·30	0·32	0·38	0·42	0·45	0 47	0·50	0·62
2·6	0·13	0·15	0·16	0·18	0·19	0·21	0·26	0·29	0·32	0·34	0·36	0·47
2·7	0·07	0·08	0·09	0·11	0·12	0·13	0·17	0·20	0·22	0·34	0·26	0·35
2·8	0·03	0·04	0·05	0·06	0·07	0·08	0·11	0·13	0·15	0·16	0·18	0·26
2·9	0·01	0·02	0·03	0·03	0·04	0·05	0·07	0·09	0·10	0·11	0·12	0·19
3·0	0·00	0·01	0·01	0·02	0·02	0·03	0·04	0·05	0·07	0·07	0·08	0·13
3·1	0·00	0·00	0·01	0·01	0·01	0·01	0·02	0·03	0·04	0·05	0·06	0·10
3·2	0·00	0·00	0·00	0·00	0·00	0·01	0·01	0·02	0·03	0·03	0·04	0·07
3·3	0·00	0·00	0·00	0·00	0·00	0·00	0·01	0·01	0·02	0·02	0·02	0·05
3·4	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·01	0·01	0·01	0·02	0·04
3·5	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·01	0·01	0·01	0·02
3·6	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·01	0·02
3·7	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·00	0·01

$t_4(-C) = 100\% - t_4(C)$

The corresponding estimator for the case of unknown variance has been given by Kolmogorov [2]

$$t_4 = \mathbf{E}(x_1 < L \mid \bar{x}, s) = I_t\left(\frac{n-2}{2}, \frac{n-2}{2}\right).$$

Here $I_t(q, q)$ denotes the incomplete Beta function:

$$I_t(q,q) = \begin{cases} 0 & \text{if} \quad t < 0 \\ \dfrac{\Gamma(2q)}{[\Gamma(q)]^2} \displaystyle\int_0^t [u(1-u)]^q \, du & \text{if} \quad 0 \leq t \leq 1 \\ 1 & \text{if} \quad t > 1 \end{cases}$$

tabulated in [6], and

$$t = \frac{1}{2} - \frac{\bar{x} - L}{s} \cdot \frac{\sqrt{n}}{2(n-1)}.$$

Lieberman and Resnikov [3] gave tables listing the values of t_3 and t_4 as functions of $\dfrac{\bar{x}-L}{\sigma}$ and $\dfrac{\bar{x}-L}{s}$, respectively.

These estimators are used in the sampling plans by variables [5].

The estimation of the quantity p is a problem of reliability theory as well. Here p is called reliability; Lloyd and Lipow (p. 203 in [4]) suggest the use of t_1 and t_2 for its estimation. On the other hand, the use of biased estimators is particularly dangerous in reliability applications: the accumulated biases may cause serious errors when estimations concerning equipments consisting of a large number of components are computed*.

The range of the mentioned tables of t_4 in [3] is restricted for the purpose of the sampling plans. Table I of this paper extends this range. The values have been computed by the same method. The accuracy of Table I is believed to satisfy practical demands in reliability applications.

Although unbiasedness is an advantageous feature of the estimators, it is not sure that the use of t_3 and t_4 is the most advantageous possibility in every case.

Easy calculation shows that

$$\mathbf{E}(t_1) = \Phi\left(\frac{L-\mu}{\sigma}\sqrt{\frac{n}{n+1}}\right),$$

i.e. the bias of t_1 is positive if $p < 1/2$; it is negative if $p > 1/2$; in the case $p = 1/2$ it is unbiased. Probably t_2 has the same property.

Lieberman and Resnikov, (p. 9 in [6]) made some numerical comparisons between the mean square errors of t_2 and t_4 and found that their relative merit depended on the value of the parameter p at which the mean square error was computed.

* The use of the unbiased estimators t_3 and t_4 is suggested by Gnedenko, Belyaev and Solovev (pp. 215, 217. in [1]).

Let us consider the case $n = 1$ if the variance is known, and the case $n = 2$ if it is unknown. In the former case

$$t_3 = \begin{cases} 1 & \text{if} \quad x_1 < L \\ 0 & \text{if} \quad x_1 \geq L. \end{cases}$$

If the variance is unknown and $n = 2$ then

$$t_4 = \begin{cases} 1 & \text{if} \quad \max(x_1, x_2) < L \\ 1/2 & \text{if} \quad \min(x_1, x_2) < L \leq \max(x_1, x_2) \\ 0 & \text{if} \quad \min(x_1, x_2) \geq L. \end{cases}$$

Using the language of quality control, in these cases both estimators disregard the information surplus of sampling by variables. In addition, the estimators have high jumps in the abscissa point L (in points where one of the abscissas is L).

These examples emphasize the necessity of further investigations.

REFERENCES

1. Гнеденко Б.В., Беляев, Ю. К. и Соловьев, А.Д.: Математические методы в теории надежности. Изд. наука, Москва, 1965.
2. Колмогоров, А.Н.: Незмешенные оценки, *Известия АН СССР сер. мат.* **14**, 303–326 (1950).
3. Lieberman, G. J., Resnikov, G. J., Sampling Plans for Inspection by Variables. *Journal of the American Statistical Associa ion* **50**, 457–516 (1955).
4. Lloyd, D. K., Lipow, M., Reliability: Management, Methods and Mathematics. Prentice Hall, Englewood Cliffs, 1962.
5. Mil–Std–414 A. Sampling Procedure and Tables for Inspection by Variables, Government Printing Office, Washington, 1959.
6. Pearson, K., Tables of the Incomplete Beta-function. Cambridge University Press, Cambridge, 1937.
7. Resnikov, G. J., Liebermann, G. J., Tables of the Non-central t-Distribution. Stanford University Press, Stanford, 1957.

EXPERIMENTELLE UNTERSUCHUNG VON EINFACH-, DOPPEL- UND MEHRFACHSTICHPROBEPLÄNEN DES MIL. STD. 105

von

O. SCHÜRZ

Institut für Angewandte Mathematik und Mechanik
Deutsche Akademie der Wissenschaften zu Berlin
Berlin, Deutsche Demokratische Republik

In der DDR haben wir den Std. 105 B, Stichprobensystem für Attributprüfung, unter den Bezeichnung TGL 14 450 standardisiert.

Nach sorgfältigen Untersuchungen der verschiedensten Systeme fiel unsere Wahl auf STD 105 B, weil dieses System

1. einen zufriedenstellenden Schutz gegen die Annahme von Erzeugnissen *unbefriedigender Qualität* und die *Rückweisung befriedigender Qualität* gibt, ohne den Hersteller untunlich zu belasten;
2. eine den höchstens Ansprüchen der Technik entsprechende Feinfühligkeit und Anpassungsfähigkeit aufweist;
3. sich verhältnismäßig einfach und wirtschaftlich handhaben läßt und
4. der Forderung gerecht wird , daß die Annahmeprüfungen den Hersteller veranlassen bzw. zwingen, die Qualität seiner Erzeugnisse kontinuiert zu verbessern.

Vom Hersteller absolut fehlerfreie Erzeugnisse erzwingen zu wollen, würde den Gestehungspreis ungünstig beeinflussen, insbesondere wenn man die Fehlerklassen, kritische Fehler, Nebenfehler und unwesentliche Fehler in einen Topf werfen würde.

Aus diesem Grunde unterteilt Std. 105 den sog. AQL-Wert in 14 Stufen von 0,015 bis 10,0. AQL, (acceptable quality level), bei uns Gutlage genannt, ist der maximale Prozentsatz fehlerhafter Einheiten oder die maximale Anzahl der Fehler je 100 Einheiten, der, zwecks Stichprobenprüfung als ein befriedigender Prozessmittelwert für jeweils eine oder mehrere bestimmte Fehlerklassen betrachtet werden kann.

In Abb. 1a ist für die dargestellten vier Stichprobenpläne der AQL-Wert = 2,5% bzw. 2,5 Fehler/100 Einh. Sie erkennen gleichzeitig, von rechts nach links gesehen, die zunehmende Trennschärfe der Pläne, wenn der Stichprobenumfang wächst. Weiter geht aus Abb. 1a die gleitende Tendenz der Risiko- oder α-Werte hervor.

In Abb. 1b, unten, wird dagegen die Verschärfung eines Planes bei gleichem Stichprobenumfang durch Verkleinerung des AQL-Wertes demonstriert.

Dem *Hersteller* wird also ein kleiner aber vernünftiger, d.h. für den Abnehmer tragbarer Spielraum an Fehlerhaftigkeit gewährt, den er unter Berücksichtigung seiner eigenen Qualitätserwartungen und derjenigen der Endverbraucher billigen darf.

Der Zweck meiner Ausführungen ist, einige Untersuchungen bezüglich der Standards 105 B und 105 D zu kommentieren.

Für die Versuche benutzten wir unter anderem ein Perlenmodell. Es besteht aus einem Gehäuse, das hier 10 000 Perlen von je 4 mm ∅ enthält. Es ist so konstruiert, daß eine gründliche Durchmischung verbürgt ist. Die Fehlerhaftigkeiten werden durch verschiedenfarbige Perlen dargestellt, hier eingeteilt in 0,5; 1,0; 2,0; 4,0 und 8,0%.

Darstellung a) der Wirksamkeit und b) der verschärften Prüfung beim Standard 105B

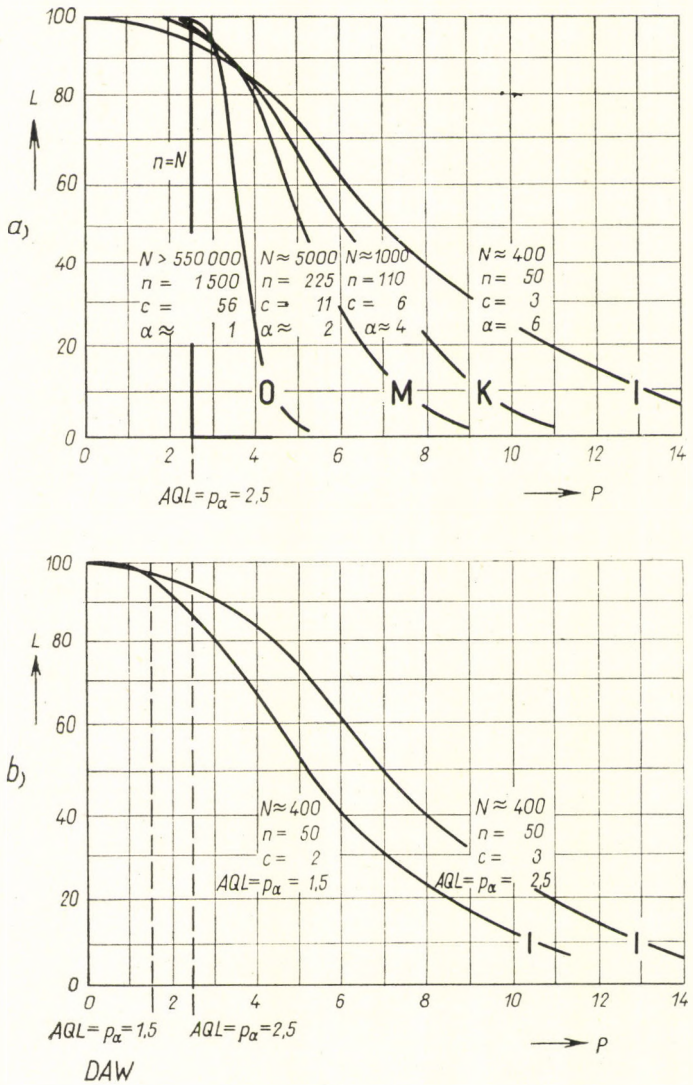


Abb. 1

Die Frontplatte enthält 300 schräggebohrte Löcher, in denen sich die Perlen fangen. Sie ist durch eine Plexiglasscheibe nach vorn abgedichtet, sodaß die gefangenen Perlen leicht abgezählt werden können.

Jede Versuchsreihe bestand aus 100 Ziehungen. Es wurden auf diese Weise mehr als 8000 Stichproben geprüft.

Der Nachteile von Modellversuchen waren wir uns bewußt. Am Modell kann man die Wirklichkeit der Betriebspraxis nicht vollständig nachahmen, bei der mit Krankheit, Urlaub u. Nachlässigkeit des Personals, mit variablen Merkmalen des Materials, verschiedenen Maschinen und Fertigungsmethoden usw. gerechnet werden muß.

Die Darstellung der Versuchsergebnisse wurde wie folgt eingeteilt: Auf Abb. 2 und 3 sieht man von oben nach unten:

1. die Rückweisungen in %;
2. den Durchschlupf AOQ (average outgoing quality), d.h. den mittleren Prozentsatz Fehlerhaftigkeit, den das Lager auf die Dauer enthalten müßte;
3. den reinen Stichprobenaufwand in Einheiten ohne Auslese der zurückgewiesenen Lose und
4. den Gesamtprüfaufwand in Einheiten einschließlich Auslese aller zurückgewiesenen Lose.

Für die Durchführung der Versuche waren zwei grundlegende Überlegungen maßgebend:

1. Die Fertigung weist einen festen Fehlerprozentsatz $\bar{p}$ auf, aber der AQL-Wert variiert.

Dieser Fall liegt vor, solange die Vertragspartner in Verhandlungen über die Höhe der AQL-Werte stehen. Die Partner fragen, mit welchen Prüfaufwänden und Folgerungen ist bei geforderten, variablen AQL-Werten zu rechnen? Also: $\bar{p}$ wird festgehalten, AQL ist variabel. Das zeigt Abb. 2.

2. Die zweite Überlegung lautet: AQL steht beim Abnehmer fest und $\bar{p}$ ist variabel. Das ist der Fall, wenn der Liefervertrag in Tätigkeit getreten ist. Die Frage lautet: Welche Prüfaufwände fallen an, welche Folgerungen sind zu ziehen, wenn bei einem festgehaltenen AQL-Wert während der Dauer des Vertrages die Qualität, d.h. der $\bar{p}$-Wert variiert?

Also:

AQL wird festgehalten, $\bar{p}$ ist variabel. Das zeigt Abb. 3.

Auf einige Kostenarten, die bei den Partnern unterschiedlich anfallen, mußte verzichtet werden, um die Anzahl der Variablen nicht unziemlich zu erhöhen. Solche Kosten sind z.B. Vertragsstrafen bei Rückweisungen, Transport- u. Korrespondenzkosten.

Bezüglich der verschärften u. reduzierten Prüfung wurde bei den Untersuchungen streng nach den Benutzungsvorschriften gearbeitet.

Wir wollen jetzt Abb. 2 ($\bar{p}$ festgehalten und AQL variabel) betrachten. Wir denken uns, der *Hersteller* bietet ein bestimmtes $\bar{p}$ an, und der Abnehmer versucht dem $\bar{p}$-Wert einen ihm zusagenden AQL-Wert anzugleichen, den er hinsichtlich der Qualität des Enderzeugnisses verantworten kann.

Es wurden zwei kontrastreiche Losumfänge gewählt, links, auf Abb. 2, ein kleiner mit $N = 500$ mit einem $\bar{p} = 2{,}8\%$ und rechts ein großer mit $N = 10\,000$ und $\bar{p} = 1{,}0\%$.

Betrachten wir zuerst die linke Seite, für $N = 500$ und $\bar{p} = 2{,}8\%$.

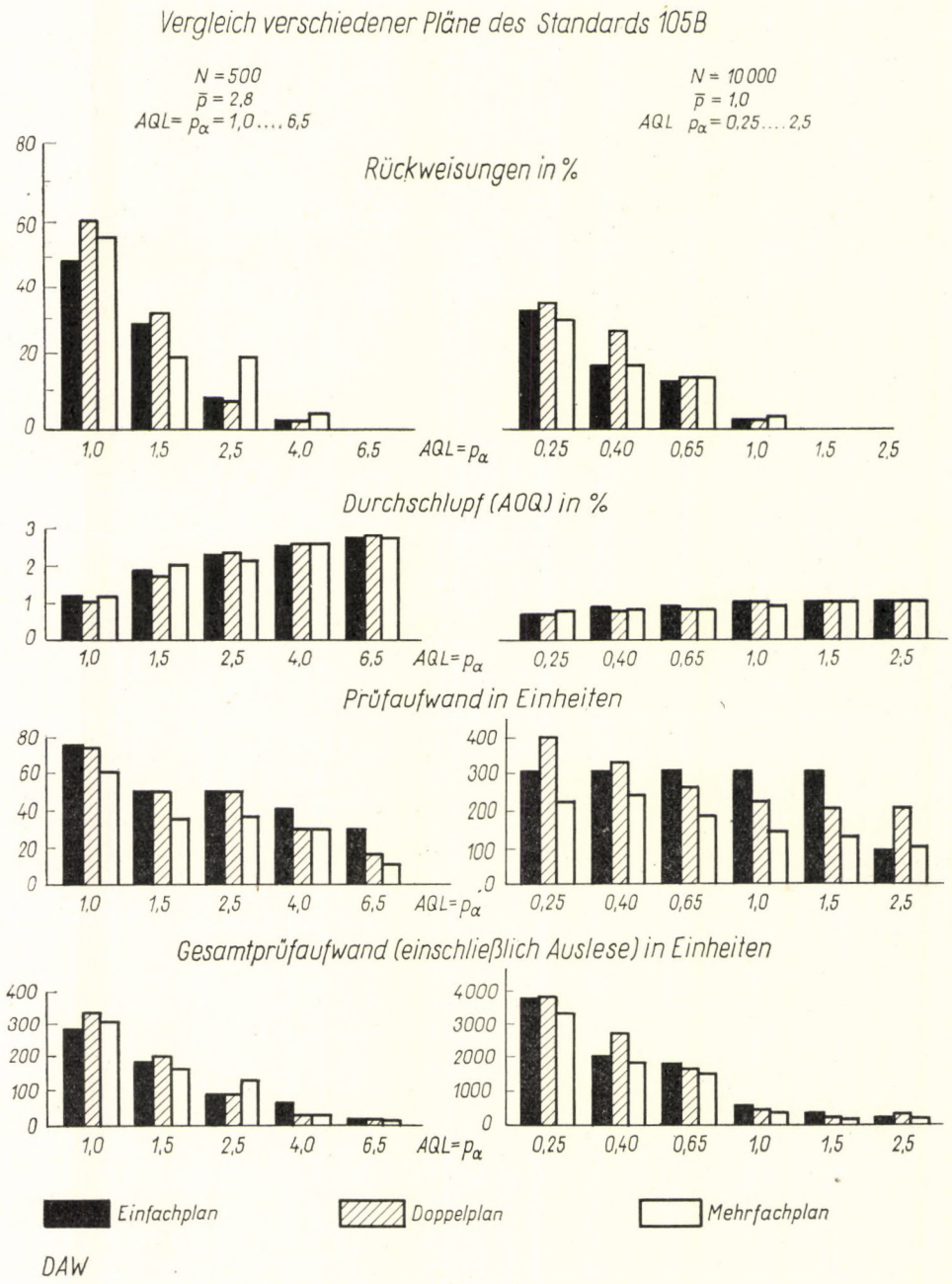

Vergleich verschiedener Pläne des Standards 105B
N = 500
p̄ = 2,8
AQL = pα = 1,0 6,5
N = 10 000
p̄ = 1,0
AQL pα = 0,25 2,5
Rückweisungen in %
80
60
40
20
0
1,0 1,5 2,5 4,0 6,5 AQL = pα 0,25 0,40 0,65 1,0 1,5 2,5
Durchschlupf (AOQ) in %
3
2
1
0
1,0 1,5 2,5 4,0 6,5 AQL = pα 0,25 0,40 0,65 1,0 1,5 2,5
Prüfaufwand in Einheiten
80
60
40
20
0
400
300
200
100
0
1,0 1,5 2,5 4,0 6,5 AQL = pα 0,25 0,40 0,65 1,0 1,5 2,5
Gesamtprüfaufwand (einschließlich Auslese) in Einheiten
400
300
200
100
0
4000
3000
2000
1000
0
1,0 1,5 2,5 4,0 6,5 AQL = pα 0,25 0,40 0,65 1,0 1,5 2,5
Einfachplan
Doppelplan
Mehrfachplan
DAW

Abb. 2

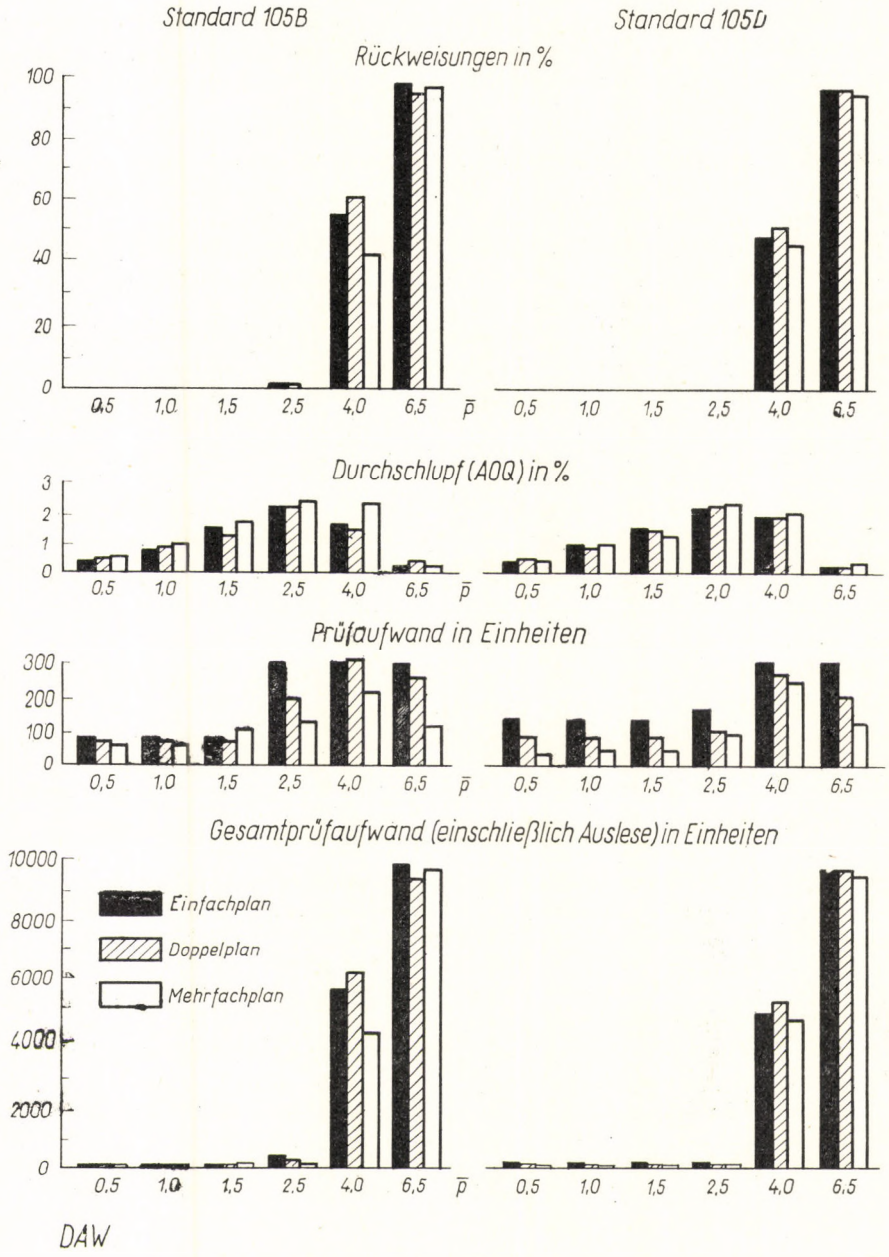

Vergleich zwischen Standard 105B und 105D für N=10000, AQL=p_α=2,5 und $\bar{p}$=0,5....6,5
Standard 105B
Standard 105D
Rückweisungen in %
100
80
60
40
20
0
0,5
1,0
1,5
2,5
4,0
6,5
$\bar{p}$
Durchschlupf (AOQ) in %
3
2
1
0
Prüfaufwand in Einheiten
300
200
100
Gesamtprüfaufwand (einschließlich Auslese) in Einheiten
10000
8000
6000
4000
2000
Einfachplan
Doppelplan
Mehrfachplan
DAW

Abb. 3

Liefert der Hersteller qualitätsgerecht, d.h. ist $\bar{p} \approx$ AQL, hier $\bar{p} = 2{,}8\%$ und AQL $= 2{,}5\%$, so werden nicht mehr als 6% Rückweisungen anfallen. Es kommt also nur die Normalprüfung in Frage.

Je kleiner wir aber die AQL wählen, desto schärfer fällt die Prüfung aus. Es werden bei AQL $= 2{,}5\%$ etwa 6%, aber bei AQL $= 1{,}5\%$ bereits 30% der Lose abgelehnt werden. Bei AQL $= 1\%$ würden es schon 60% werden.

Unter den Rückweisungen ist der Durchschlupf AOQ (average outgoing quality) dargestellt. Mit AQL $= 2{,}5\%$ geht AOQ auf etwa 2,25% zurück, bei AQL $= 1{,}0\%$ bleibt AOQ bei etwa 1,0% was aber mit etwa 60% Auslesearbeit bezahlt werden muß.

Weiter sehen wir darunter, auf den reinen Stichprobenaufwand bezogen, daß dieser z.B. bei AQL $= 1\%$ rund 13%, aber (ganz unten), auf den Gesamtprüfaufwand bezogen, etwa 60% beträgt.

Der Aufwand für AQL $= 1{,}5\%$ und $1{,}0\%$ steigt (links) deshalb so steil an, weil für kleine Lose bzw. kleine Stichproben und bei kleinen AQL-Werten keine Annahmezahlen c vorhanden sind. Man muß dann entsprechend höhere Schlüsselbuchstaben, d.h. größere Stichprobenumfänge benutzen. In extremen Fällen kann es sogar vorkommen, daß dadurch der Stichprobenumfang den Losumfang erreicht.

Betrachten wir nun die rechte Hälfte Abb. 2 für $N = 10\,000$ und $\bar{p} = 1{,}0\%$.

Es fällt auf, daß hier die Säulen kaum halb so hoch sind wie links. Dies demonstriert den Unterschied in der Wirksamkeit bei Prüfen großer Lose gegenüber kleinen Losen.

Der relative Prüfaufwand bei $N = 10\,000$ beträgt nur etwa 1/3 gegenüber dem bei Losen mit $N = 500$.

Zusammenfassend erkennen wir aus dieser Studie: Die Prüfarbeit ist reibungslos und am wirtschaftlichsten, wenn sich der Hersteller bemüht, aufgrund der Informationen der eigenen p-Karten oder den Qualitätsgeschichten des Abnehmers $\bar{p} =$ AQL oder kleiner zu liefern.

Sobald der Liefervertrag, in dem die AQL festgelegt wurde, in Tätigkeit getreten ist, sind andere Überlegungen anzustellen, nämlich: was geschieht, wenn jetzt z.B. bei festgehaltenem AQL-Wert $\bar{p}$ variiert?

Auf Abb. 3 sind die Ergebnisse von zwei Versuchsreihen mit wieder jeweils 100 Losen bei gleichen Versuchsbedingungen gezeigt.

In beiden Fällen handelt es sich um $N = 10\,000$.

Festgehalten wurde AQL $= 2{,}5\%$, während $\bar{p}$ von 0,5 . . . 6,5% variiert.

Links stehen die Ergebnisse der Versuche mit Std 105 B und rechts die von Std 105 D.

Als wir nämlich bei uns die TGL 14 450 unter Zugrundelegung des Std's 105 B unter Dach gebracht hatten, wurden wir dank der Hilfe des Prof. Dr. Stange, Aachen, vom ABC-Standard 105 D unterrichtet.

Durch die auf Abb. 3 dargestellten Versuche haben wir die beiden Systeme miteinander verglichen.

Betrachten wir zuerst die linke Seite, die Ergebnisse mit Std. 105 B.

Bei den Rückweisungen (oben), erkennen wir wieder die große Feinfühligkeit und Anpassungsfähigkeit des Systems 105. Mit $\bar{p} = 4{,}0\%$ weisen wir bereits über die Hälfte der Lose zurück und mit 6,5% praktisch alle.

Darunter sehen wir wieder den Durchschlupf AOQ (average outgoing quality). Er geht bei AQL $= \bar{p}$, hier 2,5% auf etwa 2,3% zurück und wird bei kleiner, wie auch bei größer werdendem $\bar{p}$ merkbar kleiner.

In der III. Abteilung, darunter, ist wieder der reine Stichprobenaufwand aufgezeichnet. Für $\bar{p} = 1{,}5$ 1,0 u.o., 5% ist er wegen der dann durchzuführenden reduzierten Prüfung auf 1/3 des Aufwandes für $\bar{p} = 2{,}5\%$ heruntergegangen; für $\bar{p} >$ AQL bleibt er gleich.

Ganz unten schnellt der Gesamtprüfaufwand bereits bei $\bar{p} = 4\%$ auf etwa 50% hinauf u. erreicht bei $\bar{p} = 6{,}5\%$ fast 100%, d.h. es müssen praktisch alle Lose ausgelesen werden.

Auf der rechten Seite der Abb. 4 sind die Versuchsergebnisse mit Standard 105 D, dem ABC-System dargestellt*.

Im großen und ganzen gesehen unterscheiden sie sich von 105 B nicht besonders.

Die Doppelpläne von 105 D sind durchweg weniger aufwendig als die Einfachpläne, während es bei Std. 105 B, hier z.B. bei $\bar{p} = 4{,}0\%$ vorkommt, daß die Doppelpläne teuerer als die Einfachpläne arbeiten.

Ein auffälliger Unterschied ist aber bei $\bar{p} =$ AQL $= 2{,}5\%$ zu erkennen: Dort ist der Aufwand der reinen Stichprobenprüfung kleiner als bei 105 B, da bei 105 D die reduzierte Prüfung eher durchgeführt werden darf.

Für $\bar{p} <$ AQL arbeiten die Einfachpläne bei 105 B billiger als bei 105 D, weil die Reduzierung bei 105 B 1/5, bei 105 D dagegen 2/5 der Aufwände für die Einfachpläne beträgt.

Bei den Doppel- u. Mehrfachplänen liegt 105 D deshalb etwas günstiger weil 105 D eigene Tabellen für reduzierte Prüfung besitzt.

Wenn wir aber schließlich die unterste Stufe, die Gesamtaufwände betrachten, fällt der Unterschied nicht mehr so sehr ins Gewicht.

Wenden wir uns nun den Qualitätsgeschichten zu, die in Abb.4 aufgezeichnet sind:

Dafür wurde die Versuchsreihe $N = 10\,000$, AQL $= 2{,}5\%$ mit Std.105 B von Abb. 3 zugrundegelegt.

Gemäß der Benutzungsanweisung für den Std. 105 D braucht der Prozeßdurchschnitt nicht mehr errechnet zu werden. Da wir aber zwecks Entscheidung über die Prüfart (reduziert, normal, verschärft) die Fehlerzahlen, Anzahl der Lose und gesamt geprüfte Stücke auf alle Fälle registrieren und addieren müssen, entfällt eigentlich nichts weiter als eine einzige Division, d.h. eine Rechenschiebereinstellung.

Ich habe, um den Arbeitaufwand bei der Qualitätsgeschichte von 105 B mit den Aufschreibeaufwendungen bei 105 D beurteilen und vergleichen zu können, eine umfangreiche Studie durchgeführt, die ich Interessenten gern zeigen kann.

Danach bleibt der Arbeitsaufwand eigentlich *gleich*.

Die registrierten Prüfergebnisse bei 105 D sind aber bei weitem nicht so übersichtlich wie bei den Qual.-Geschichten von 105 B, wo wir mit einem Blick den Qualitätsverlauf über lange Zeiträume erfassen können. Außerdem brauchen die Buchhaltungen die Fehler-Prozentwerte für ihre Kalkulationen.

* ABC=America-Britain-Canada

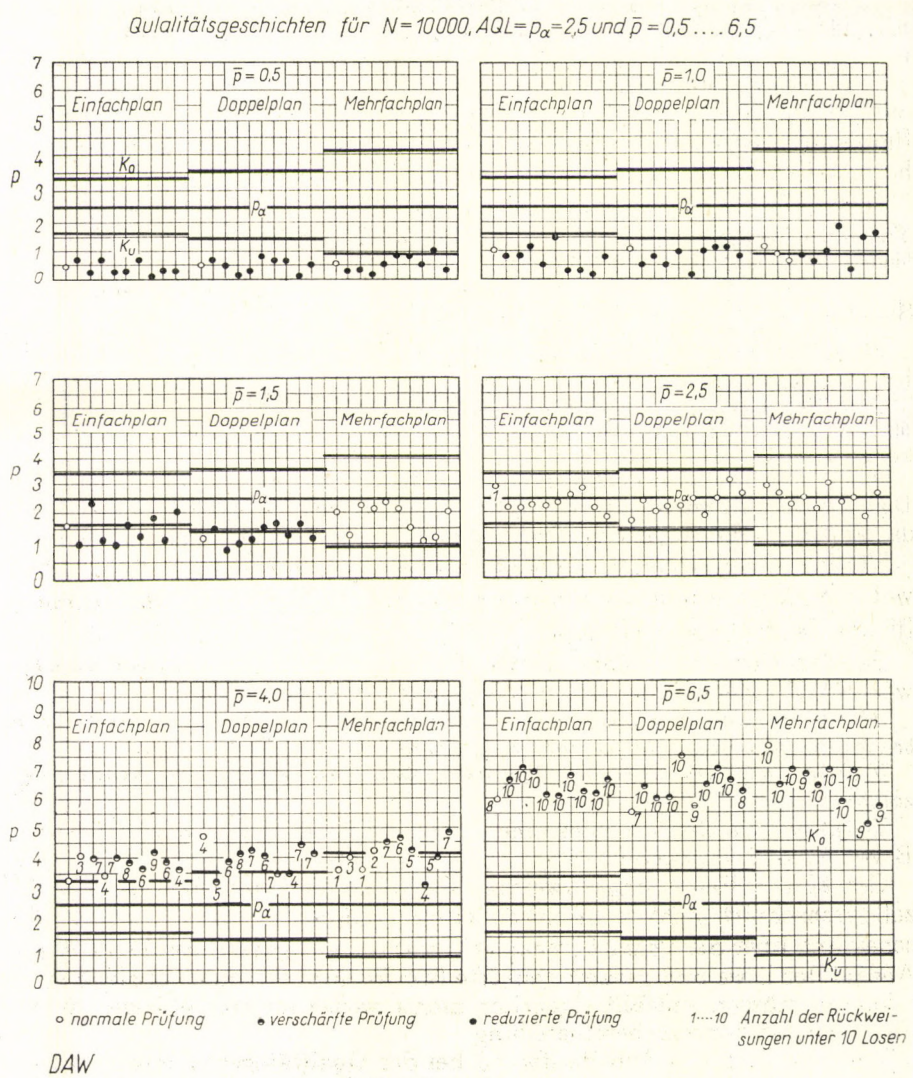


ABB. 4

Die Qualitätsgeschichte, die ja weiter nichts ist als eine p-Karte, hat aber den großen Vorteil, daß der Abnehmer den Hersteller viel leichter kontrollieren u. ihm überzeugender Hinweise geben kann, wo er seine Hebel zur Verbesserung ansetzen muß. Kurz, die Qualitätsgeschichte ist und bleibt ein Hauptbuch für die Qualität und eine fruchtbare Diskussionsgrundlage für die Vertragspartner.

Sie sehen auf Abb. 4, daß der AQL-Wert überall derselbe ist. Sie sehen weiter, daß die Qualitätsgeschichten

1. nach $\bar{p}$ von 0,5 ... 6,5% und
2. nach den Prüfarten, Einfach-Doppel- u. Mehrfachplänen eingeteilt wurden.

Die Zahlen bedeuten die jeweilige Anzahl der Rückweisungen bei 10 aufeinanderfolgenden Losen. Leere Kreise zeigen an, daß die betr. Serie von 10 Losen normal, volle dagegen, daß reduziert geprüft wurde. Das Halbmondszeichen bedeutet verschärfte Prüfung.

Nun ist noch die Frage zu beantworten, inwieweit sich die Einfachpläne, die Doppelpläne und die Mehrfachpläne unterscheiden.

Aus den Qualitätsgeschichten sehen Sie, daß die Streuungen zwischen den Prüfarten nicht sehr unterschiedlich sind.

Betrachten wir Abb. 3, wo die Einfachpläne schwarz, die Doppelpläne schraffiert u. die Mehrfachpläne leer angelegt sind, so erkennen wir, daß die Doppelpläne bei 105 B nicht immer, wie zu erwarten, billiger arbeiten als die Einfachpläne.

Das ist ein Nachteil des Std. 105 B gegenüber 105 D.

Bei Std. 105 D sind die Doppelpläne hinsichtlich der reinen Stichprobenprüfung immer billiger. Das liegt darin begründet, daß bei 105 D der Umfang der 2. Stichprobe gleich dem der ersten ist, während bei Std. 105 B die zweite Stichprobe doppelt so groß wie die erste ist.

Es ist noch eine letzte Frage zu klären:

Erzielen wir wesentliche Einsparungen dadurch, daß wir die erste Stichprobe bei Erreichen der Rückweiszahl abbrechen?

Das ist nämlich bei Std. 105 D erlaubt. Man könnte diesen Passus natürlich auch bei 105 B zulassen, das wäre eine Frage der Vereinbarung zwischen den Partnern.

Solange $\bar{p} \leq$ AQL, hier wieder $\leq 2{,}5\%$ gibt es keine Diskussion, sobald aber $\bar{p}$ den AQL-Wert wesentlich überschreitet, geht natürlich der Prüfaufwand der reinen Stichprobenarbeit zurück.

Wenn Sie aber nochmals Abb. 3. vergleichen wollen, so werden Sie erkennen, daß bereits bei $\bar{p} = 4{,}0\%$ mehr als 50% der Lose beanstandet werden und bei $\bar{p} = 6{,}5\%$ brauchten wir gar nicht erst mit Stichproben zu beginnen.

Bitte legen Sie sich die Frage vor, ob Sie als Abnehmer einen Zustand und wielange dulden würden, bei dem mehr als 30% der erwarteten Lose abgelehnt, zurückgesandt oder von Ihnen selbst ausgelesen werden müßten — das wäre wohl untragbar, auch wenn der Lieferer zu jeweils einer Vertragsstrafe von 6% verpfiichtet wäre.

LITERATUR

1. Mil. Std. 105D, Sampling Procedures and Tables for Inspection by Attributes, U. S. Dept. of Def., Washington, D. C., April 29, 1963.
2. Hamaker, H. C., Report on Sampling Committee of the European Organization for Quality Control 1959.
3. Stephens, K. S., Larson, K. E., An Evaluation of the Mil-Std-105D System of Sampling Plans. *Industrial Quality Control*, Vol. **23.** Nr. 7 Januar 1967.
4. TGL 14 450 (DBR-STD.) Stichprobenpläne für die Attributprüfung. Buchhaus Leipzig, Abt. Stds.

ADDITIVE PROZESSE MIT ZUFALLSSPRÜNGEN IN DEN ABSORPTIONSWALL UND LEBENSDAUERMODELLE BEI DEN BRUCHERSCHEINUNGEN IN STARREN KÖRPERN WELCHE DIESEN PROZESSEN ENTSPRECHEN

von

J. SEDLÁČEK

Staatliches Forschungsinstitut für Wärmetechnik

Prag, Tschechoslowakei

1. EINLEITUNG

Die statistische Theorie des Mechanismus von Brüchen bei Metallen, die in diesem Artikel kurz beschrieben wird, soll zuerst mit einer kurzgefaßten historischen Übersicht der statistischen Ermündungstheorie eingeleitet werden, aus der sich die entsprechende Theorie entwickelt hat.

Vor dem 2. Weltkriege erschien im Jahre 1938 in der Fachliteratur eine umfangreiche experimentelle Studie von Emil Ravilly über die Ermüdung unter dem Titel «Contributions à l'étude de la rupture de fils métalliques soumis à des torsions alternées«, in der zum ersten Male auf eine bedeutende Streuung in den Werten der Ermüdungsdauer N hingewiesen wurde, falls Ermüdungsversuche unter denselben Bedingungen wiederholt wurden. (Es handelte sich um eine Ermüdung bei wechselnder Torsionsbeanspruchung). Die Ergebnisse dieser Arbeit sind jedoch damals ohne jede weitere mathematische Auswertung geblieben, da zu jener Zeit keine Theorie vorhanden war, die die Gründe der großen Streuung in den Werten der Bruchleben erklären können hätte und eine mathematische Verarbeitung sowie eine mathematische Beschreibung der Ermüdung gestattet hätte. Zu jener Zeit hat man auch keinen geeigneten mathematischen Ausdruck für die Wöhlersche Ermüdungskurve eingeführt die der technischen Fachwelt in einer überwiegenden Mehrzahl von Fällen in der Form einer eingeführt »nach Augenmaß« geführten Linie $W(x = \log N,\ y = \sigma)$ vorgelegt wurde, die durch die Punkte von Koordinatenpaaren (x_j, y_j), die man auf experimentellem Wege gewonnen hat, hindurchging.

Die Arbeit von Ravilly hat indirekt darauf hingewiesen, daß man die bisherige klassische Art der Beschreibung der Ermündung mit der Wőhlerschen Kurve irgendwie genauer formulieren muß, da klar nachgewiesen wurde, daß die Ermüdung eine Zufallserscheinung ist und infolgedessen eine auf die mathematische Statistik sich gründende Lösung erfordert. Den ersten Schritt im Bestreben um eine genauere Formulierung der Ermüdung hat der schwedische Physiker W. Weibull getan, der mit Hilfe der mathematischen Statistik zu einer exakteren Beschreibung der Ermüdungseigenschaften der Metalle in der Form von Ermüdungskurven $(\log N, \sigma)p$ =konst. durch die Gleichungen

$$\log N_P = k_P - m_P \log (\sigma - \sigma_{c,P}) , \tag{1}$$

$$(0 \leq P \leq 1) ,$$

gelangte, so daß die frühere Beschreibung mit Hilfe der Wöhlerschen Kurve um eine neue Größe — die Gewähr P für das Versagen — auf die Ermüdungslebensdauer N bei einer gegebenen konstanten Amplitude der Wechselspannungen σ erweitert wurde.

Diesen Schritt Weibulls, der in (L 2) beschrieben wird, kann man als einen weiteren Beitrag zu einer neuen Auffassung der Ermüdung und als Grundlage der statistischen Theorie der Ermüdung von Metallen überhaupt betrachten. Der Klarheit halber sei noch hinzugefügt, daß die Gleichung (1) von Weibull nicht theoretisch begründet, sondern intuitiv abgeleitet wurde. Dabei hat Weibull jedoch die damals als neu geltende Theorie der größenordnungsmäßigen Stichprobenkoeffizienten völlig ausgenutzt. Die Erprobung der Eignung der Gleichung (1) für die Ausnützung von größenordnungsmäßigen Stichprobenkoeffizienten wurde dann auf Grund der experimentellen Ergebnisse von Ravilly (s. L. 1) vorgenommen.

Um die gesamte statistische Ermüdungstheorie übersehen zu können, charakterisiert man diesen ersten Beitrag von Weibull zur statistischen Ermüdungstheorie durch die Beziehung

$$\left.\begin{matrix}\sigma = \text{houst.} \\ \theta = \text{const.}\end{matrix}\right\} \Rightarrow (\log N, P), \tag{2}$$

(σ = Amplitude, θ = Periode der Zyklen, N = Zahl der Zyklen, P = Wahrscheinlichkeit)

Den zweiten Schritt in der statistischen Ermüdungstheorie haben A. M. Freudenthal und E. J. Gumbell in ihrer gemeinsamen Arbeit (aus den Jahren 1951 und 1953) »On statistical Interpretation of Fatigue Tests« — (*Proceedings of the Royal Statistical Society*, Series A. **216**, 1953) getan. Diese waren bestrebt, die Lücke — eigentlich eine Begründung des Wahrscheinlichkeitsmodells der Lebensdauer — in der Weibullschen Theorie auszufüllen. Ihre Lösung hat, genauer gesagt, bezweckt, die Dichtefunktion der Zufallsveränderlichen N (die die Zyklenzahl der Schwingungsspannung bis zum Bruch angibt), welche die Dauer des Ermüdungslebens bezeichnet, theoretisch zu begründen oder die Funktion $l\,(N)$ zu bestimmen, welche die elementare Wahrscheinlichkeit

$$l(N_\sigma)\,dN = Pr\{N \leq \text{ Ermüdungsdauer } \leq N + dN\} \tag{3}$$

angibt.

In der Übersicht der statistischen Ermüdungstheorie sei dieser zweite Schritt durch die symbolische Beziehung

$$\left.\begin{matrix}\sigma = \text{houst.} \\ \theta = \text{const.}\end{matrix}\right\} \Rightarrow l(N)_\sigma \tag{4}$$

gekennzeichnet.

Den dritten Schritt, der nach der Veröffentlichung einer Reihe Arbeiten von A. M. Freudenthal erfolgte, kann man einer Reihe von Verfassern (z. B. Epstein, Z. W. Birnbaum, S. C. Saunders, S. V. Serensen, V. P. Kogajev) zuschreiben. Er knüpft an die theoretischen Ergebnisse der Arbeiten von

Freudenthal an, die er dadurch vertiefte, daß die Ermüdung als ein spezieller Fall der allgemeinen Lebensdauerprobleme betrachtet wurde. Diese Theoretiker waren bestrebt die Lebensdauermodelle aus dem Urwesen der Brucherscheinungen – aus dem Mechanismus der Brüche in den Metallen – abzuleiten. Diesen dritten Schritt in der statistischen Ermüdungstheorie kann man als einen Übergang von den Ermüdungsbrüchen zu Brüchen aller möglichen Typen bei Metallen bezeichnen und diesen mit der Beziehung

$$\left.\begin{matrix}\sigma = \text{houst.}\\ \theta = \text{const.}\end{matrix}\right\} \Rightarrow \mu_z(t) = -\frac{S'(t)}{S(t)} \Rightarrow 1 - L_z(t) \Rightarrow l(t)\,, \tag{5}$$

$$\text{resp.} \quad z(t) \Rightarrow \mu_z(t) \Rightarrow T(z) \Rightarrow l(t)\,.$$

charakterisieren, was in Worten gesagt folgendes bedeutet:

Der regelmäßige zyklische Prozess der Wechselspannungen ruft in der Struktur des Metalls irgendeinen Beschädigungsvorgang hervor, der am Ende durch einen Bruch in der Zeit τ abgeschlossen ist, wobei $\mu_z(t)$ die Intensität des Versagens angibt. τ ist dabei eine zufällige Größe, die eine beliebige Funktiondichte $l(t)$ und Verteilungsfunktion $L(t)$ besitzt. Die für die Funktion $L(t)$ sich ergebende Grundbeziehung hat die Form

$$L(t) = 1 - \mathrm{e}^{-\int_0^t \mu(x)dx}\,.$$

Als ein Beispiel der Lösung (5) kann vor allem die Arbeit von Z. W. Birnbaum und S. C. Saunders (s. *L* 5) gelten, die unter dem Titel »A Statistical Model for Life-Length of Materials« erschienen ist.

2. GRUNDLAGEN DER ALLGEMEINEN STATISTISCHEN THEORIE VON BRUCHERSCHEINUNGEN BEI METALLEN

Die erwähnte Arbeit ist uns zum Hauptausgangspunkt für die Formulierung der theoretischen Hauptaufgabe geworden, die auf die Bildung einer stochastischen Theorie der Brüche bei Metallen hinzielt. Man kann sie kurz durch die symbolische Beziehung

$$X(t) \Rightarrow Z(t) \Rightarrow T(t)_{z_{\text{krit}}} \Rightarrow l(t) \tag{6}$$

darstellen, was in Worten gesagt das folgende ausdrückt:

(a) Der primäre Prozeß der Brucherscheinungen, der z. B. in den früheren Arbeiten auf dem Fachgebiet der statistischen Ermüdungstheorie als ein regelmäßiger Prozeß der Wechselspannungen von konstanter Amplitude und Periode betrachtet wird, wird hier allgemein als irgend ein stochastischer Prozeß der Zeitwirkung äußerer Kräfte $X(t)$ betrachtet. Dabei ist der Ermüdungsfall bloß ein spezieller Fall, in dem der Prozeß $X(t)$ durch einen stochastischen zyklischen Spannungsprozess vorgestellt wird, in dem allgemein sowohl die Amplitude σ als auch die Periode θ zufällige Größen sein können.

(b) Der Prozeß $Z(t)$, der an den stochastischen Prozess $X(t)$ anknüpft, ist ein stochastischer Prozess der Akkumulation von Teil-Sprungbeschädigungen von (zufälliger) Größe ξ (oder ein Zufallsprozess der Anhäufung von Störungen).

(c) Der Prozeß $T(t)_{z\text{krit}}$ ist dann ein Abgangsprozess, der dem betrachteten Mechanismus der Brucherscheinungen entspricht, der durch den stochastischen Prozeß $Z(t)$ beschrieben ist. Die Funktion $l(t)$ definiert dann die Verteilung von Lebensdauern und kann aus dem Prozeß $Z(t)$ durch eine Inversion abgeleitet werden.

Zum Unterschied von der erwähnten Arbeit von Birnbaum und Saunders wird also in vorliegenden Artikel eine allgemeinere Theorie vorgeschlagen, die sich von den bisherigen Theorien in den folgenden Punkten unterscheidet:

1. Der primäre Prozeß der Brucherscheinungen $X(t)$, der der Zeitwirkung der äußeren Kräfte (der Belastung von Konstruktionen) entspricht, ist im allgemeinen ein stochastischer Prozeß.

2. Die Erläuterung aller Brucharten gründet sich auf eine stochastische Erläuterung des Prozesses der Anhäufung von Beschädigungen, der durch den Prozeß $X(t)$ hervorgerufen ist, wobei die Konstruktion des Modells des Prozesses $Z(t)$ an die Komponenten des Prozesses $X(t)$ anknüpft, d.i. von der tatsächlichen Natur des Mechanismus der Brucherscheinungen ausgeht. Dabei wird eine wesentliche Vertiefung des Prozesses $Z(t)$ dadurch vorgenommen, daß man die Möglichkeit eines heftigen Sprunges im Prozeß $Z(t)$ zuläßt, wodurch eine Möglichkeit gegeben ist, sowohl die spröden Gewaltbrüche als auch die zufälligen Betriebsbrüche, die den Charakter einer Havarie aufweisen (Brüche durch eine heftige Überlastung), zu erläutern.

3. ALLGEMEINES MODELL DES PROZESSES DER BRUCHENTWICKLUNG $Z(t)$

3.1. *Kurze Charakteristik des theoretischen Modells des Prozesses* $Z(t)$

Man betrachte den stochastischen Prozeß $Z(t)$ und setze dabei voraus:

(a) Der Parameter t ist stetig und nimmt die Werte $0 \leq t < \infty$ an.

(b) Die Veränderliche ξ, die die Zustände im Prozeß $Z(t)$ definiert, nimmt die Werte aus dem Intervall $0 \leq \xi < Z_{\text{krit}} < \infty$ an.

(c) Die Zustandsveränderungen im Prozeß $Z(t)$ können von zweierlei Art sein:

Zustandsveränderung vom Typus (1), die im beliebigen Zeitintervall $(t, t+\Delta t)$ eintreten kann und die durch die Beziehung

$$Z_{t+\Delta t} = Z_t + \xi \tag{7}$$

definiert ist; in diesem Ausdruck ist ξ eine nichtnegative Zufallsveränderliche, die durch die Wahrscheinlichkeitsdichte $\gamma(x)$ definiert ist.

Zustandsveränderung vom Typus (2), die gleichfalls in beliebigem Zeitintervall $(t, t+\Delta t)$ unter jedem beliebigen Zustand $Z_t \in (0, Z_{\text{krit}})$ eintreten kann und die durch die Beziehung

$$Z_{t+\Delta t} = Z_t + (Z_{\text{krit}} - Z_t) \tag{8}$$

definiert ist; dies stellt einen plötzlichen raschen Übersprung aus dem Zustand $Z = Z_t$ in der Zeit t in den Zustand Z_{krit} im Verlauf des Zeitintervalls von Länge Δt vor;

(d) Der Prozeß $Z(t)$ wird in der Zufallszeit $\tau \in (0, \infty)$ im bestimmten festen Zustand $Z_{\text{krit}} \leq \infty$ absorbiert.

(e) Für die Wahrscheinlichkeit p_1 des Zustandekommens einer Veränderung vom Typus (1) gilt die Beziehung

$$p_1 = \lambda_1(t)\,\Delta t + 0(\Delta t)\,, \tag{9}$$

($0(\Delta t)$ bestimmt die Wahrscheinlichkeit von 2 und mehr Veränderungen), indem das Zustandekommen einer Veränderung vom Typus (2) durch die Wahrscheinlichkeit

$$p_2 = \lambda(z, t)\,\Delta t + 0(\Delta t) = \lambda_2(z) + \lambda_3(t) + 0(\Delta t) \tag{10}$$

gegeben ist.

Abschließend sei noch erwähnt, daß die Veränderliche $Z(t)$ in den Realisationen dieses theoretischen Modells des Prozesses $Z(t)$ den Umfang der Beschädigungen in der Struktur der Metalle bestimmt, die durch die zeitwirkenden Kräfte hervorgerufen worden sind und daß aus den Beziehungen (7) bis (10) folgt, daß es sich um einen bestimmten Typus der additiven stochastischen Prozesse handelt, in denen die allgemeinere Abhängigkeit der Brucherscheinung, und zwar nicht nur vom Zeitverlauf des Alterns, sondern auch von der Größe des zerstörten Teiles, durch die Beziehungen (8) und (10) respektiert wird.

Der plötzliche heftige Sprung in den kritischen Zustand kann dann als typisch für die meisten Brüche bei Metallen betrachtet werden.

3.2 *Mathematisch-statistischer Ausdruck des Störungsprozesses $Z(t)$ und des entsprechenden inversen Prozesses $T(z)$*

Das Modell des Prozesses der Bruchentwickelung bei Metallen, das durch den Prozess $Z(t)$ im Abschnitt 3.1 definiert ist, kann am besten durch diejenige Funktion dargestellt werden, die die Dauer des Prozesses ausdrückt. Man bestimme also, mit welcher Wahrscheinlichkeit keine andere Veränderung als diejenige vom Typus (1) im Zustand $z < z_{\text{krit}}$ im Verlauf eines beliebigen Zeitintervalls $(t, t + \Delta t)$ eintreten wird. Diese Wahrscheinlichkeit wird also bestimmen, daß

$$\text{Pst}\,\{z \leq Z_t < z + dz, \tau > t\} = p_0(z, t)\,dz, \tag{11}$$

$$(0 \leq t < \infty\,,\ 0 \leq z \leq z_{\text{krit}})\,.$$

Daraus ergibt sich weiter, daß

$$\int_0^z p_0(\xi, t)\,d\xi = \text{Pst}\,\{\xi \leq z, \tau > t\}\,,$$

und folglich

$$1 - \int_0^z p_0(\xi, t)\,d\xi = \text{Pst}\,\{\xi \leq z, \tau \leq t\} = L_z(t)\,. \tag{12}$$

Die Beziehung (12) kann man in Worten als Wahrscheinlichkeit des Austritts aus den Zuständen $\xi_1 \in 0, Z)$ im Prozeß $Z(t)$ in die Zustände $\xi_2 \in \in< Z, Z_{krit})$ und zwar in der Zufallszeit $\tau \in (0, t)$ interpretieren, so daß man die Beziehung (12) als einen Ausdruck des Prozesses $T(z)$ betrachten kann, der durch eine Inversion des Prozesses gewonnen wurde.

Aus der Voraussetzung der Stetigkeit der Größe τ ergibt sich dann, daß die Wahrscheinlichkeitsdichte $l_z(t)$, mit deren Hilfe die elementare simultane Wahrscheinlichkeit

$$\mathrm{Pst}\{\xi \leq z, t \leq \tau < t + dt\} = l_z(t)\, dt\,, \tag{13}$$

ausgedrückt werden kann, durch die Ableitung

$$l_z(t) = \frac{d}{dt}[L_z(t)] = -\int_0^z \left\{\frac{\partial}{\partial t}[p_0(\xi, t)]\right\} d\xi \tag{14}$$

definiert wird.

Abgekürzt schreibt man also:

$Z(t)$ ist ausgedrückt durch die Funktion $p_0(Z, t)$,

$$T(z) \text{ ist ausgedrückt durch die Funktion } 1 - \int_0^z p_0(\xi, t)\, d\xi. \tag{15}$$

Das Lebensdauermodell wird durch die Dichtefunktion definiert:

$$l_z(t) = -\int_0^z \left\{\frac{\partial}{\partial t}[p_0(\xi, t)]\right\} d\xi. \tag{16}$$

3.3. *Grundgleichungen der Prozesse $Z(t)$ und $T(z)$*

Wir werden uns an dieser Stelle nicht mit einer ausführlichen Ableitung der Grundgleichung des Prozesses $Z(t)$ befassen; darauf beziehen sich die Arbeiten (L 10) und (L 15). Es sei nur erwähnt, daß es sich um Gleichungen vom Kolmogoroff–Fellerschen Typus handelt und zwar in der Form

$$\frac{\partial p_0(z, t)}{\partial t} = \lambda(t)\,[\gamma(z) * p_0(z, t)] - p_0(z, t)\,[\lambda_1(t) + \lambda(t, z)], \tag{17}$$

die in einem speziellen einfachen Fall, in dem im Prozeß bloß Veränderungen vom Typus (2) zustandekommen, die Form

$$\frac{\partial p_0(z, t)}{\partial t} = \lambda_1(t)\,\{[\gamma(z) * p_c(z, t)] - p_c(z, t)\} \tag{18}$$

annimmt. Dabei stellt der Ausdruck $[\gamma(z) * p_0(Z, t)]$ in diesen beiden Formeln eine Konvolution der Funktionen γ und p_0 vor. Die Lösung dieser Integro-Differentialgleichungen kann unter bestimmten begrenzenden Voraus-

setzungen für die Funktion $\lambda(z, t)$ mit Hilfe der Laplaceschen Transformation vorgenommen werden, wodurch die Gleichungen (17) und (18) dann in eine partielle Differentialgleichung überführt werden. Die Lösung ist in den Arbeiten (L 10) und (L 15) ausführlich durchgeführt.

Für die Gleichung (18) sei erwähnt, daß dieselbe, falls die Laplacesche transformierte Funktionen $p_0(z, t)$ und $\gamma(z)$ mit Bezug auf z als $P_0(s, t)$ bzw. $g(s)$ bezeichnet wird, in eine einfache Gleichung

$$\frac{\partial P_0(s,t)}{\partial t} = \lambda_1(t)\, P_0(s,t)\,[g(s) - 1]\,, \tag{19}$$

übergeht, die bei den Anfangsbedingungen

(a) $P_0(s, 0) = 0$ für alle $s \geq 0$,

(b) $P_0(0, t) \in < 0{,}1 >$,

eine Lösung in der Form

$$P_0(s,t) = \mathrm{e}^{-\Lambda(t)}\, \mathrm{e}^{g(s)\Lambda(t)} \tag{20}$$

bietet, wobei

$$\Lambda(t) = \int_0^t \lambda(x)\, dx\,,$$

so daß der Prozeß $Z(t)$ durch die Funktion $p_0(z, t)$ in der Form

$$p_0(z,t) = e^{-\Lambda(t)}\, \Lambda\{\mathrm{e}^{g(s)\Lambda(t)}\}\,, \tag{21}$$

ausgedrückt wird, wo das Symbol Λ das abzubildende Original bezeichnet.

Der inverse Prozeß $T(z)$ ist dann auf Grund der Beziehung (12) und der bekannten Formel

$$\mathscr{L}\left\{\int_0^z \varphi(x)\, dx\right\} = \frac{1}{s}\,\mathscr{L}\{\varphi(x)\} = \frac{1}{s}\,\Phi(s) \tag{22}$$

durch die Beziehung

$$\mathscr{L}\{L_z(t)\} = \frac{1}{s}\{1 - P_0(s,t)\} = \frac{1}{s}\{1 - \mathrm{e}^{-\Lambda(t)[1-g(s)]}\}\,, \tag{23}$$

ausgedrückt, so daß die Dichtefunktion $l_z(t)$ durch die Laplacesche transformierte Funktion

$$\mathscr{L}\{l_z(t)\} = \frac{1 - g(s)}{s} \cdot \frac{d\Lambda(t)}{dt}\, \mathrm{e}^{-\Lambda(t)[1-g(s)]} \tag{24}$$

bestimmt wird.

Damit ist ein allgemeiner Ausdruck sowohl für den Prozeß der Anhäufung der Beschädigungen als auch der entsprechenden Lebensdauermodelle gegeben und es hängt nur von der speziellen Form der Funktionen $\gamma(Z)$ und $\lambda(t)$ ab.

4. SPEZIELLE FÄLLE VON BRÜCHEN BEI METALLEN

4.1. *Plastischer Bruch*

Bei dem plastischen Bruch kann man den Prozeß $X(t)$ z. B. in der Form eines stochastischen Prozesses der kurzfristigen Stoßbelastungen von gleicher Größe R_ε betrachten (Abb. 1):

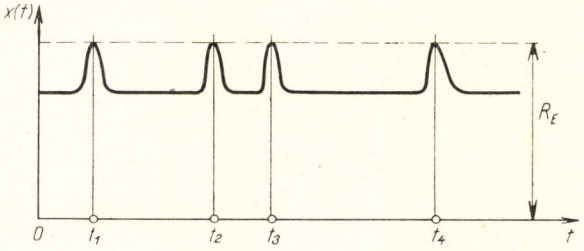


ABB. 1

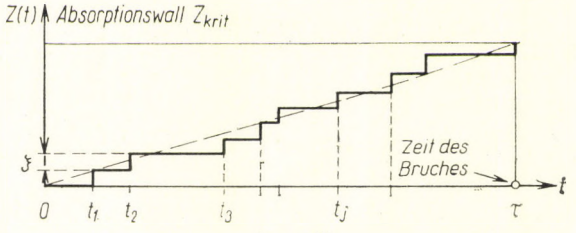


ABB. 2

Diesem primären Prozeß $X(t)$ kann man nun den Prozeß der Anhäufung von Beschädigungen $Z(t)$ mit zeitzufällig realisierten Teilbeschädigungen von Konstanter Größe ($\xi = C \pm \varepsilon$, $\varepsilon \approx 0$) beiordnen, was in Abb. 2 schematisch dargestellt ist.

Wird nun die vorige Theorie auf diesen Mechanismus des Bruches angewendet, so ergibt sich daraus ein Lebensdauermodell mit einer Wahrscheinlichkeitsdichte $l(t)$ von der Form

$$l(t) = \frac{\lambda^{k+1}}{\Gamma(k+1)} t^k \mathrm{e}^{-\lambda t}, \quad \left(k = \left[\frac{z_{\mathrm{krit}}}{c}\right]\right), \tag{25}$$

(die Pearsonsche Kurve von Typ III).

Ein allgemeineres Modell des plastischen Bruches gründet sich dann auf die Voraussetzung, die Teilzuwächse ξ im Beschädigungsprozess $z(t)$ seien zufällig und durch die Wahrscheinlichkeitsdichte

$$\gamma(x) = ae^{-ax}, \tag{26}$$
$$(a > 0, \; x \in (0, \infty)),$$

bestimmt. Unter dieser Voraussetzung von Prozeß $Z(t)$ ergibt sich für die Dichtefunktion der Zufallsveränderlichen τ der Ausdruck

$$l_z(t) = \lambda e^{-\lambda t}[e^{-az} + a\lambda t z e^{-az} + \ldots]. \tag{27}$$

$$\left(t \in (0, \infty), 0 \leqq Z \leqq Z_{krit} < \infty\right).$$

4.2 *Klassischer Ermüdungsbruch*

Als Beispiel eines Ermüdungsbruchs sei ein einfacher Fall gewählt, in dem der Prozeß $X(t)$ ein regelmäßiger, zyklischer Prozeß der Wechselspannungen $\sigma(t)$ ist, was der Wöhlerschen klassischen Ermüdung (Abb. 3) entspricht.

Während dieses primären Prozesses $X(t)$ setzt sich im Prozeß $Z(t)$ vor allem der Faktor der »Alterung« durch, der durch die Intensität $\lambda(t)$ definiert ist. Wählt man für $\lambda(t)$ die an den früheren Ermüdungsmodellen erprobte Beziehung in der Form

$$\lambda(t) = ct^{\alpha}, \tag{28}$$

$$(c > 0, \ \alpha \geq 1)$$

und setzt man von den Teilzuwächsen ξ_j, $(j = 1, 2, 3, \ldots)$ voraus, sie seien praktisch gleich, so daß für sie

$$\xi_j = c \lg \sigma_a \pm \varepsilon, \ (j = 1, 2, \ldots) \tag{29}$$

$$(\varepsilon \to 0)$$

gilt, so ergibt sich aus dem stochastischen Prozeß $Z(t)$ ein Lebensdauermodell $l(t)$, das durch die Wahrscheinlichkeitsdichte in der Form

$$l(t) = \frac{\lambda c^K}{\Gamma(K)} e^{-ct^{\alpha}} t^{\alpha K - 1}, \tag{30}$$

$$\left(c > 0, \alpha \geq 1, K \gg 0, t \in (0, \infty)\right)$$

ausgedrückt ist.

Durch die Beziehung (30) ist eine breite Klasse der Verteilungen, wie z. B. Gamma-, χ^2, Weibullsche Verteilung usw. definiert. Diesen Fall kann man gleichfalls als eine neue theoretische Begründung und Ableitung des Wei-

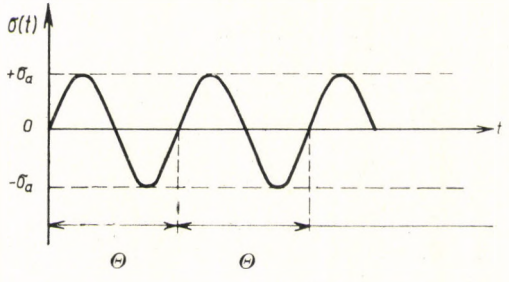


Abb. 3

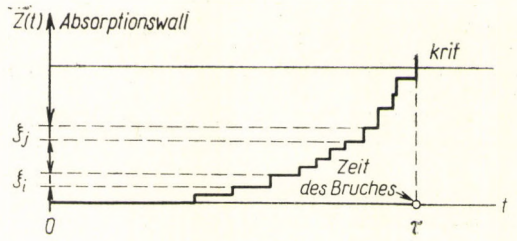


ABB. 4

bullschen Modells zur Interpretation der Wöhlerschen Ermüdung betrachten. Der Prozeß $Z(t)$ der zu diesem Fall des Bruches gehört, ist in schematischer Form in Abb. 4 wiedergegeben.

4.3 *Betriebliche spröde Gewaltbrüche*

Man wähle für diesen Typus der Brüche den Prozeß $X(t)$ in derartiger Form, daß man zwar die statische Belastung $X(t) = q(t) = q_{st} = \text{const}$, betrachtet, aber das Arbeitsdiagramm in der Zeit zufällige Stoßspitzen wirkender Kräfte von Zufallsgrößen q_i aufweist (Abb. 5):

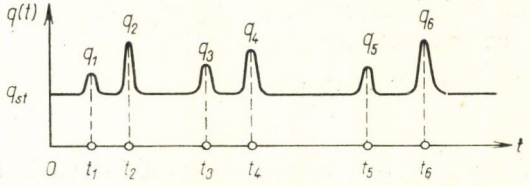


ABB. 5

Eine nähere Spezialisierung des Prozesses $X(t)$ ist dadurch vorgenommen, daß man für die Spitzen des Prozesses q_i eine Exponentialverteilung in Betracht zieht, so daß die Dichtefunktion $h(q)$ der Wahrscheinlichkeitsverteilung durch

$$h(q) = a\,\mathrm{e}^{-aq}, \tag{31}$$

$$(a > 0,\ q \in (0, \infty))$$

gegeben ist. Zum Unterschied von den beiden vorigen Realisationen, die man durch die Gleichung (18) ausdrücken kann, ist der Typus der Brüche von allgemeinerer Natur und soll sich auf die Erläuterung gründen, die auf der Gleichung (17) basiert. In diesem Falle ist die Verteilungsfunktion $L(t)$ der Zufallsveränderlichen τ durch einen Komplizierteren Ausdruck

$$L(t) = 1 - \exp\left[-\lambda_1 t - \int_0^t \lambda_3(t)dt + \frac{a\lambda_1}{a\lambda_2}\int_0^{\lambda_2 t} \frac{\mathrm{e}^{-\beta x}}{x + \frac{a}{\alpha}}\,dx\right]$$

$$(\lambda_2(z) = \lambda_2 z,\ \lambda_2 \neq 0)\,, \tag{32}$$

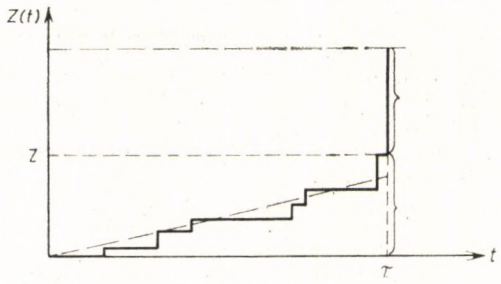


ABB. 6

gegeben, den man in der Beziehung zur Intensität der Sterblichkeit weiter analysieren kann. Dies wird aber in diesem Artikel nicht bezweckt. Die entsprechende Realisation des Prozesses $Z(t)$ ist für diesen Fall in Abb. 6 wiedergegeben.

5. SCHLUß

Zum Schluß der Erläuterung dieser neueren Theorie der Brucherscheinungen bei Metallen sei noch in Kurzem diese Bemerkung erwähnt:

In den letzten fünf Jahren kann man in der physikalischen Erforschung des Mechanismus der Brüche vor allem eine große Initiative in der Erforschung der spröden Gewaltbrüche beobachten. Diese Forschungsarbeiten sind fast ausschließlich auf eine deterministische Weise gelöst ebensowie deren Anknüpfung an das endliche Ziel, auf das diese Forschung hinzielt, d. i. die Bewertung der Zeitstandfestigkeit oder der Lebensdauer bei einem bestimmten Mechanismus der Ausbreitung und Anhäufung von Störungen der Kohäsion der Metalle ist nicht gesichert. Die stochastische Theorie der Brüche versucht, diesen Nachteil zu beseitigen, d. i. sie führt eine Grundlage ein, die eine Verbindung von Ergebnissen der physikalischen Erforschung des Mechanismus der Brüche mit einer richtigen Interpretation der Festigkeit und der Lebensdauer gestattet und zwar nicht nur im Falle von Proben, sondern auch von mechanischen Systemen und ganzen Metallkonstruktionen. Eine mathematisch-statistische Lösung dieser Aufgabe soll auch weiterhin fortgesetzt werden und dieser Artikel beschreibt bloß einen der Schritte der theoretischen Lösung der komplexen Auffassung der Festigkeit von Metallen und starren Körpern.

LITERATUR

1. RAVILLY, É., Contributions à l'étude de la rupture de fils métalliques soumis à des torsions alternées. Publications Scientifiques et Techniques du Ministère de l'Air (1938).
2. WEIBULL, W., A Statistical Representation of Fatigue Failure. *Transaction of the Royal Inst. Techn. Stockholm*, No. **27**, (1949).
3. FREUDENTHAL, A. M., GUMBEL, E. J., On the Statistical Interpretation of Fatigue Test. *Proceedings of the Royal Society, Series A*, **216**, (1953).
4. CRAMER, H., Collective Risk Theory. The Jubilee Volume Skandia Insurance Company 1955.

5. Birnbaum, Z. W., Saunders, S. C., A Statistical Model for Life-length of Materials. *JASA*, **55**, March (1958).
6. Mercer, A., Smith, C. S., A Random Walk in which the Steps Occur Randomly in Time. *Biometrika*, Part **1**, **2**, June (1959).
7. Sedláček, J., Režný, Z., Untersuchung der Ermüdungseigenschaften der Werkstoffe für Wälzlager mit Hilfe der Verfahren der mathematischen Statistik. *Forschungsbericht SVÚTT*-59–01014.
8. Sedláček, J., Režný, Z., Einführung in die statistische Ähnlichkeitstheorie bei der Lösung der Festigkeitsprobleme. *Forschungsbericht SVÚTT*–60–01016.
9. Sedláček, J., Statistische Interpretation der physikalischen und mechanischen Eigenschaften der Metalle und Erzeugnisse. *KAMS-Konferenz*. Prag, 15.–17. Mai 1962.
10. Sedláček, J., Stochastische Modelle des Prozesses der Entwicklung von Störungen in der Struktur der Metalle. *Forschungsbericht SVÚTT*–62–01028.
11. Sedláček, J., Dauerfestigkeit bei stochastischen Schwingungen linearer mechanischer Systeme. *Forschungsbericht SVÚTT*–63–01019.
12. Němec, J., Untersuchungen über die Ausbreitung von Ermüdungsrissen in den Körpern. *Strojírenství*, Nr. 5. **12**, (1962).
13. Seremsen, S. V., Kogajev, V. P., Statistische Gesetzmässigkeiten der Ähnlichkeit bei Ermüdungsbrüchen. *Strojírenství*, Nr. 9. **13**, 1963.
14. Frost, N. E., The Propagation of Fatigue Craks in Various Sheet Materials. *National Engineering Laboratory Report, AB, Div.* No *50/58*.
15. Sedláček, J., Additive Prozesse mit Zufallssprüngen in den Absorptionswall und ihre Anwendung bei der Interpretation der Brüche in den Metallen. (Übergegeben der Zeitschrift *Aplikace matematiky)*.

ON THE MEAN NUMBER AND SIZE OF OPAQUE PARTICLES IN TRANSPARENT BODIES

by

Z. ŠIDÁK

Mathematical Institute of the Czechoslovak Academy of Sciences
Prague, Czechoslovakia

1. STATEMENT OF THE PROBLEM

In the analysis of microstructure of metals and alloys we often study cases in which some particles (commonly assumed to be sphere shaped) are dispersed in the metal, and we wish to find their number and size. Usually we have at our disposal an opaque sample of the metal, so that we observe particles only on a plane section of the sample, and measure only the diameters of particles cut by that section; these diameters are, of course, less or equal to the actual diameters of the whole particles. Similar situations occur frequently also in biology, etc., and there is a number of publications devoted to the study of this problem, see e.g. [1] — [6] with further references.

Recently I was asked to solve the following modified, more general problem. We have a section of some thickness τ, and we observe it, or better to say look through it, in a microscope. The basic medium of the section appears in the microscope transparent, while the particles are opaque. Thus we observe the maximal diameter of that part of a particle which is contained in the section.

The problem may be mathematically formulated e.g. in the following way.

Suppose that ν points are randomly dispersed (i.e. independently and with the uniform distribution) in a cube which has the length of its edge b. Around each point as a centre a sphere is circumscribed having the diameter D, possibly different for different spheres. Denote for convenience also λ_V the number of particles per unit volume, i.e. $\lambda_V = \nu/b^3$.

The number of particles ν may be supposed constant or random. In the latter case the defining relation for λ_V should be naturally $\lambda_V = \mathbf{E}\,\nu/b^3$ where, as usual, the letter $\mathbf{E}$ denotes the expectation of a random variable and will be used in this sense throughout the whole paper. From the practical point of view and for our further development the question, whether ν is considered to be constant or random, plays no essential role, and we shall suppose it constant. On the other hand, it is usually necessary to consider the diameters D of particles as random variables, and we shall also do so; moreover, suppose that the diameters of different particles are independent and have the same density f. Nevertheless, in proofs it brings some advantage to take first D fixed and then employ a conditional argument.

Continuing the description of our model, let us have a section of thickness τ cut out from the middle of the cube parallel to its sides. Let A be the event

that some fixed particle will be cut or contained in this section. The number of particles (random) cut by the section or contained in it will be denoted by N, and their mean number per unit area by λ_A, i.e. $\lambda_A = \mathbf{E}N/b^2$. (Here we must take the area instead of the volume since we are looking through the section and observing, in fact, only a plane picture.) The observed diameter of a particle, that is the maximal diameter of that part of a particle which is contained in the section, will be denoted by C and its density by g. Naturally, we must bear in mind that considering the observed diameter C makes sense only under the condition that A has occurred so that the density g is, in fact, a conditional density $g(. \mid A)$.

For practical applicability let us also assume that ν is large and the size of the particles is small enough so that a sufficient number of particles can be observed in the section in order to obtain reasonable estimates, but that the thickness is still much smaller so that it is practically impossible for a particle to lie directly above the other in the section and hide it in the resulting microscopic picture.

Obviously, for $\tau = 0$ we come back to the known problem concerning plane sections, and we shall see that setting $\tau = 0$ in our subsequent formulae reduces them to the known ones. However, we shall also see that the case $\tau > 0$ presents more difficulties owing to the impossibility of a simple estimation of $\mathbf{E}D$.

2. THE NUMBER N OF OBSERVED PARTICLES

First, let us find the probability $\mathbf{P}(A)$ that some fixed particle will be observed in the section. For later convenience, define an auxiliary random variable Z equal to the distance of the centre of the particle from the middle plane of the section (see Fig. 1). Clearly, our basic assumption on random position of particles implies that Z has the uniform distribution in the interval $[0, 1/2\ b]$, and that it is independent of D. Supposing the condition $D = x$ is satisfied, the event A occurs if and only if the centre of the particle is placed either in the section or within the distance $\frac{x}{2}$ from the section, i.e. if and only if $0 \leqq Z < \frac{1}{2}\tau + \frac{1}{2}x$. Thus

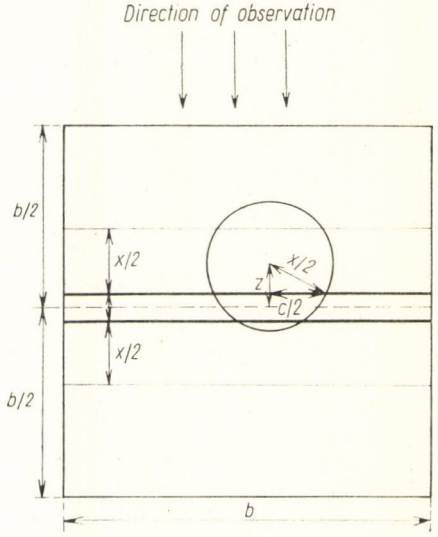


Fig 1.

$$\mathbf{P}(A \mid D = x) = \frac{\frac{1}{2}\tau + \frac{1}{2}x}{\frac{1}{2}b} = \frac{\tau + x}{b},$$

whence

$$\mathbf{P}(A) = (\tau + \mathbf{E}D)/b. \qquad (1)$$

Defining auxiliary random variables $N_i (i = 1, \ldots, \nu)$ by $N_i = 1$ if the ith particle is observed in the section and $N_i = 0$ otherwise, we obtain

$$\mathbf{E}N = \mathbf{E}\Big(\sum_{i=1}^{\nu} N_i\Big) = \sum_{i=1}^{\nu} \mathbf{E}N_i = \nu(\tau + \mathbf{E}D)/b\,.$$

It follows

$$\lambda_A = \mathbf{E}N/b^2 = \nu(\tau + \mathbf{E}D)/b^3 = \lambda_V(\tau + \mathbf{E}D)\,,$$

so that the actual overall density λ_V of particles may be estimated by means of the equation

$$\lambda_V = \lambda_A/(\tau + \mathbf{E}D) \tag{2}$$

provided $\mathbf{E}D$ is known.

3. THE OBSERVED DIAMETER C OF A PARTICLE

The general density of the diameter D of an arbitrary particle has been denoted by f. However, we shall need now the conditional density $f(. \mid A)$ of the diameter D under the condition A that the particle is observed in the section. By the classical Bayes rule

$$f(x \mid A) = \frac{\mathbf{P}(A \mid D = x)f(x)}{\int_0^\infty \mathbf{P}(A \mid D = x)f(x)dx} = \frac{(x + \tau)b^{-1}f(x)}{\int_0^\infty (x + \tau)\, b^{-1}f(x)dx} = \frac{(x + \tau)f(x)}{\mathbf{E}D + \tau} \tag{3}$$

For finding the density $g(c) = g(c \mid A)$ we start from the identity

$$\mathbf{P}(C < c \mid A) = \int_0^c \mathbf{P}(C < c \mid A, D = x)f(x \mid A)dx +$$

$$+ \int_c^\infty \mathbf{P}(C < c \mid A, D = x)f(x \mid A)dx\,. \tag{4}$$

Since $C \leqq D$, the probability in the first integral is equal to 1. Further, since under the condition A and $D = x$ the variable Z has the uniform distribution in $\left[0, \frac{1}{2}x + \frac{1}{2}\tau\right]$, the probability in the second integral becomes

$$\mathbf{P}(C < c \mid A, D = x) = \mathbf{P}\left(\sqrt{\left(\frac{x^2}{4} - \frac{c^2}{4}\right)} + \frac{\tau}{2} < Z \leqq \frac{x}{2} + \frac{\tau}{2} \,\middle|\, A, D = x\right) =$$

$$= \frac{x - \sqrt{(x^2 - c^2)}}{x + \tau}\,.$$

Combining the last results with (3) and (4) we find

$$\mathbf{P}(C < c \mid A) = \int_0^c \frac{(x+\tau)f(x)}{\mathbf{E}D + \tau}\,dx + \int_c^\infty \frac{x - \sqrt{(x^2 - c^2)}}{x + \tau} \cdot \frac{(x + \tau)f(x)}{\mathbf{E}D + \tau}\,dx\,. \tag{5}$$

Differentiation then gives

$$g(c) = \frac{(c+\tau)f(c)}{\mathbf{E}D+\tau} - \frac{cf(c)}{\mathbf{E}D+\tau} + \int_c^\infty \frac{c}{\sqrt{(x^2-c^2)}} \cdot \frac{f(x)}{\mathbf{E}D+\tau}\,dx = \tag{6}$$

$$= \frac{\tau f(c)}{\mathbf{E}D+\tau} + \frac{c}{\mathbf{E}D+\tau}\int_c^\infty \frac{f(x)}{\sqrt{(x^2-c^2)}}\,dx\,.$$

The kth moment of C is

$$\mathbf{E}C^k = \frac{\tau}{\mathbf{E}D+\tau}\int_0^\infty c^k f(c)\,dc + \frac{1}{\mathbf{E}D+\tau}\int_0^\infty c^{k+1}\int_c^\infty \frac{f(x)}{\sqrt{(x^2-c^2)}}\,dx\,dc.$$

Changing the order of integration in the last integral and subsituting $c = x\,t^{1/2}$ (as in the known case with $\tau = 0$) leads to

$$\mathbf{E}C^k = \frac{1}{\mathbf{E}D+\tau}\left[\tau\mathbf{E}D^k + \frac{1}{2}\,\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}(k+2)\right)}{\Gamma\left(\frac{1}{2}(k+3)\right)}\,\mathbf{E}D^{k+1}\right]. \tag{7}$$

In particular

$$\mathbf{E}C^{-1} = \frac{1}{\mathbf{E}D+\tau}\left[\tau\mathbf{E}D^{-1} + \frac{\pi}{2}\right], \tag{8}$$

$$\mathbf{E}C = \frac{1}{\mathbf{E}D+\tau}\left[\tau\mathbf{E}D + \frac{\pi}{4}\,\mathbf{E}D^2\right]. \tag{9}$$

Denoting by S the surface of the particle, and using (9), we find the mean surface

$$\mathbf{E}S = \mathbf{E}\left\{4\pi\left(\frac{1}{2}D\right)^2\right\} = \pi\mathbf{E}D^2 = 4\mathbf{E}C(\mathbf{E}D+\tau) - 4\tau\mathbf{E}D\,, \tag{10}$$

and the mean total surface of all particles per unit volume (specific surface)

$$\lambda_V\mathbf{E}S = 4\,\mathbf{E}C\,.\;\lambda_A - 4\tau\mathbf{E}D\,\frac{\lambda_A}{\tau+\mathbf{E}D}\,. \tag{11}$$

4. DETERMINATION OF $\mathbf{E}D$

An important characteristic of the particles is $\mathbf{E}D$, and, moreover, if we knew it, we should be able to estimate λ_V, $\mathbf{E}D^2$, $\mathbf{E}S$ and the mean total surface per unit volume on the basis of Equations (2), (9), (10) and (11), re-

spectively. For the case $\tau = 0$, previously treated in literature, $\mathbf{E}D$ can be estimated on the basis of (8) which yields in this case $\mathbf{E}D = \frac{\pi}{2} \cdot \frac{1}{\mathbf{E}C^{-1}}$. Unfortunately, in our case $\tau > 0$ we lack this possibility, because (8) involves now two unknown quantities $\mathbf{E}D$ and $\mathbf{E}D^{-1}$.

We can try for example the following method for reducing the number of unknowns in (8) and (9). Setting $\psi(x) = x^{-1}$ in the inequality $\psi(\mathbf{E}\,D) \leqq \mathbf{E}\,\{\psi(D)\}$, which is known to be valid for an arbitrary convex function ψ, we get $(\mathbf{E}\,D)^{-1} \leqq \mathbf{E}D^{-1}$. Thus, using this inequality in (8) we get

$$\mathbf{E}C^{-1} \geqq \frac{1}{\mathbf{E}D + \tau}\left[\tau \frac{1}{\mathbf{E}D} + \frac{\pi}{2}\right],$$

which is equivalent to a quadratic inequality for $\mathbf{E}D$:

$$Q_1(\mathbf{E}D) = \mathbf{E}C^{-1}\,(\mathbf{E}D)^2 + \tau\mathbf{E}C^{-1}\,\mathbf{E}D - \frac{1}{2}\pi\,\mathbf{E}D - \tau \geqq 0\,. \tag{12}$$

Since $Q_1(0) = -\tau < 0$ and $\lim\limits_{u \to \pm\infty} Q_1(\mathrm{u}) = \infty$, there are two real roots u_1, u_2 of the equation $Q_1(\mathrm{u}) = 0$, and they satisfy $u_1 < 0 < u_2$. Of course, $\mathbf{E}D > 0$, so that the solution of (12) is $\mathbf{E}D \geqq u_2$, or explicitly

$$\mathbf{E}D \geqq \frac{\frac{1}{2}\pi - \tau\mathbf{E}C^{-1} + \sqrt{\left[\frac{1}{4}\pi^2 - \pi\tau\,\mathbf{E}C^{-1} + \tau^2(\mathbf{E}C^{-1})^2 + 4\tau\,\mathbf{E}C^{-1}\right]}}{2\,\mathbf{E}C^{-1}}. \tag{13}$$

A similar procedure may be performed with the relation (9). Now taking $\psi(x) = x^2$ amounts to getting the well-known Schwarz inequality $(\mathbf{E}D)^2 \leqq \mathbf{E}D^2$. After substitution into (9) and a slight simplification another quadratic inequality for $\mathbf{E}D$ is found

$$Q_2(\mathbf{E}D) = \frac{1}{4}\pi(\mathbf{E}D)^2 + \tau\mathbf{E}D - \mathbf{E}C \cdot \mathbf{E}D - \tau\mathbf{E}C \leqq 0\,. \tag{14}$$

Again $Q_2(0) = -\tau\,\mathbf{E}C < 0$ and $\lim\limits_{v \to \pm\infty} Q_2(v) = \infty$, and consequently $Q_2(v) = 0$ has again two real roots v_1, v_2, for which $v_1 < 0 < v_2$. However, $\mathbf{E}D > 0$ so that the solution of (14) is

$$0 < \mathbf{E}D \leqq \frac{\mathbf{E}C - \tau \pm \sqrt{[(\mathbf{E}C)^2 - 2\tau\,\mathbf{E}C + \tau^2 + \pi\tau\,\mathbf{E}C]}}{\frac{1}{2}\pi}. \tag{15}$$

Combining the inequalities (13) and (15) and replacing $\mathbf{E}C$ and $\mathbf{E}C^{-1}$ by their sample estimates, we obtain an "estimate" of the interval in which $\mathbf{E}D$ is lying.

Let us also remark that if

$$\mathbf{E}D \cdot \mathbf{E}D^{-1} = \frac{1}{2}\pi \tag{16}$$

the equality (8) becomes

$$\mathbf{E}C^{-1} = \frac{\pi}{2} \cdot \frac{1}{\mathbf{E}D}$$

for whichever value τ, but this is precisely the corresponding equality for the case $\tau = 0$. This means that if (16) holds, the quantity τ in (8) may be neglected. Unfortunately, only $\mathbf{E}D \cdot \mathbf{E}D^{-1} \geqq 1$ is true generally.

If it is possible to perform two groups of observations with two different thicknesses τ_1 and τ_2 of the section, the problem simplifies considerably as was kindly pointed out to me by V. Horálek. Naturally, the quantities $\mathbf{E}D$ and $\mathbf{E}D^2$ do not change with τ whereas the expectation of C does; denote these latter expectations by $\mathbf{E}_1C$ for the thickness τ_1 and $\mathbf{E}_2C$ for τ_2. Then (9) can be written

$$\mathbf{E}_1C = \frac{1}{\mathbf{E}D + \tau_1}\left[\tau_1 \mathbf{E}D + \frac{\pi}{4}\mathbf{E}D^2\right],$$

and similarly with the index 2 in place of the index 1. The solution of these two equations is

$$\mathbf{E}D = \frac{\tau_2\mathbf{E}_2C - \tau_1\mathbf{E}_1C}{\mathbf{E}_1C - \mathbf{E}_2C + \tau_2 - \tau_1}.$$

Of course, a similar procedure can be done with Equation (8).

5. LOG-NORMAL AND MAXWELL'S DISTRIBUTION OF D

If the diameter D of particles has the log-normal density

$$f(x) = \frac{1}{x\sigma\sqrt{(2\pi)}}\exp\left[-\frac{(lg\, x - lg\, \mu)^2}{2\sigma^2}\right] \quad \text{for} \quad x \geqq 0 \tag{17}$$

the problem of finding $\mathbf{E}D$ and the parameters μ, σ becomes rather easy. It is well known that in this case

$$\mathbf{E}D^k = \mu^k \exp\left[\frac{1}{2}k^2\sigma^2\right]$$

for arbitrary k, so that Equations (8) and (9) are now

$$\mathbf{E}C^{-1} = \frac{1}{\mu e^{\frac{1}{2}\sigma^2} + \tau}\left[\tau\mu^{-1}\, e^{\frac{1}{2}\sigma^2} + \frac{\pi}{2}\right], \tag{18}$$

$$\mathbf{E}C = \frac{1}{\mu e^{\frac{1}{2}\sigma^2} + \tau}\left[\tau\mu\, e^{\frac{1}{2}\sigma^2} + \frac{\pi}{4}\mu^2\, e^{2\sigma^2}\right]. \tag{19}$$

In solving practical problems, $\mathbf{E}C^{-1}$ and $\mathbf{E}C$ in (18) and (19) can be replaced by their sample estimates, and the resulting two equations can be solved for μ and σ unknown, e.g. by the following method. Introduce the new quantities $y = e^{\frac{1}{2}\sigma^2}$ and $z = \mu e^{\frac{1}{2}\sigma^2}$. Then Equations (18) and (19) become

$$\mathbf{E}C^{-1} = \frac{1}{z+\tau}\left[\tau\frac{y^2}{z} + \frac{\pi}{2}\right], \tag{20}$$

$$\mathbf{E}C = \frac{1}{z+\tau}\left[\tau z + \frac{\pi}{4}z^2y^2\right]. \tag{21}$$

From (21) we have

$$y^2 = \frac{(\mathbf{E}C - \tau)\,z + \tau\mathbf{E}C}{\frac{1}{4}\pi z^2}$$

which can be inserted into (20) giving after an easy manipulation

$$\frac{1}{4}\pi\mathbf{E}C^{-1}z^4 + \frac{1}{4}\pi\left(\tau\mathbf{E}C^{-1} - \frac{1}{2}\pi\right)z^3 - \tau(\mathbf{E}C - \tau)\,z - \tau^2\,\mathbf{E}C = 0\,.$$

This is an equation of the fourth degree for z, and it can be solved at least numerically.

An interesting situation occurs if D has Maxwell's distribution

$$f(x) = \frac{x}{\sigma}\exp\left[-\frac{x^2}{2\sigma}\right] \quad \text{for} \quad x \geqq 0\,; \tag{22}$$

observe that in this case $\mathbf{E}D = \sqrt{\left(\frac{1}{2}\pi\sigma\right)}$. S.D.Wicksell [6] has shown for the case $\tau = 0$ that whenever D has the density (22) then C has precisely the same density (22). This assertion is true even in the general case $\tau \geqq 0$ as it can be readily seen. In fact, inserting (22) into (6) we obtain the density of C

$$g(c) = \frac{1}{\mathbf{E}D + \tau}\left[\tau f(c) + c\int_c^\infty \frac{x}{\sigma\sqrt{x^2 - c^2}}e^{-\frac{x^2}{2\sigma}}dx\right].$$

Introducing in the integral a new variable $v = (x^2 - c^2)\,(2\,\sigma)^{-1}$ and an easy subsequent calculation shows $g(c) = f(c)$ which has been asserted.

6. ESTIMATION OF THE DENSITY $f(x)$ OF D

The relation (6) between the density g of C and the density f of D suggests a possibility of the approximative estimation of f. The solution of this problem for $\tau = 0$ has been described at sufficient length (including a discussion of its shortcomings, difficulties and various modifications) by J. Likeš [4], and the general case $\tau \geqq 0$ brings only a very small change into it, so that we can be here quite brief.

Assume that the diameters D can take the values from a finite interval $0 \leqq D \leqq D_m$, and consider $k + 1$ points $0, \Delta, 2\Delta, \ldots, i\Delta, \ldots, k\Delta$ in this interval where Δ is some positive number.
Denote

$$\lambda_V^{(i)} = \lambda_V \int_{(i-1)\Delta}^{i\Delta} f(x)\,dx, \quad \lambda_A^{(j)} = \lambda_A \int_{(j-1)\Delta}^{j\Delta} g(c)\,dc\,.$$

The latter quantities may be estimated from observations, and we aim to express $\lambda_V^{(i)}$ by means of $\lambda_A^{(j)}$. By (6) we have

$$\lambda_A^{(j)} = \lambda_V \left[\tau \int_{(j-1)\Delta}^{j\Delta} f(c)\,dc + \int_{(j-1)\Delta}^{j\Delta} c \int_{c}^{D_m} \frac{f(x)}{\sqrt{(x^2 - c^2)}}\,dx\,dc\right].$$

Changing the order of integration in the last integral, dividing the region of integration into sub-intervals $[(i-1)\Delta, i\Delta]$, and replacing in each such interval the integration variable x by $i\Delta$, we obtain the system of approximate equations

$$\lambda_A^{(j)} = [\tau + \Delta\sqrt{(2j-1)}]\,\lambda_V^{(j)} + \Delta \sum_{i=j+1}^{k} \left[\sqrt{\left(i^2 - (j-1)^2\right)} - \sqrt{(i^2 - j^2)}\right]\lambda_V^{(i)}$$

for $j = 1, 2, \ldots, k$. Thus considering $\tau > 0$ adds only the member $\tau\lambda_V^{(j)}$ to the system for $\tau = 0$ derived in [4]. On replacing $\lambda_A^{(j)}$ by their estimates we get a system of linear equations for $\lambda_V^{(i)}$ which is easy to solve. A similar system of equations may be obtained by another approximation, namely by replacing x by the centre of each sub-interval rather than by its upper endpoint; this possibility is also shown in [4].

REFERENCES

1. Horálek, V., Příspěvek k otázce hodnocení struktury materiálu. (A contribution to the study of the structure of materials) *Aplikace matematiky* **3**, 376–383 (1958).
2. Likeš, J., Stanovení střední velikosti částic struktury kovů a slitin. (On the determination of mean particle size in studies of structure of metals and alloys) *Aplikace matematiky* **7**, 315–323 (1962).
3. Likeš, J., On the Problem of Particle Number and Size Determination in Opaque Bodies. *Acta Technica Acad. Sci. Hung.* **42**, 325–360 (1963).
4. Likeš, J., Stanovení počtu a velikosti částic dispersní fáze metodami kvantitativní metalografie. (Determination of number and size of particles of a dispersion phase by methods of quantitative metallography.) *Rozpravy Čs. akademie věd, řada technických věd* **74**, No. 1 (1964).
5. Saltykov, S. A., Stereometric metallography. (In Russian.) Moscow 1959.
6. Wicksell, S. D., On the Corpuscle Problem. *Biometrika* **17**, 84 (1925).

DIE WIRTSCHAFTLICHSTE VERTEILUNG DES PRÜFAUFWANDS AUF DIE VERSCHIEDENEN STUFEN EINES MODELLS MIT MEHREREN ZUFALLSKOMPONENTEN

von

K. STANGE

Institut für Statistik und Wirtschaftsmathematik
der Technischen Hochschule
Aachen, Deutsche Bundesrepublik

1. DAS MODELL. GESAMTHEIT UND PROBE. AUFGABENSTELLUNG

Es liege eine Grundgesamtheit in der Form eines *Stufenmndells* oder *Schachtelmodells* vor, d. h. eine Ausgangsgesamtheit wird in der *ersten* Stufe in N_1 Einheiten unterteilt. Jede dieser N_1 *Einheiten erster Stufe* wird in N_2 *Einheiten zweiter Stufe* aufgegliedert usw. Bezeichnet man die Abweichungen der Merkmalwerte vom jeweiligen Mittelwert in den einzelnen Stufen mit

$$\alpha_i, \beta_{ij}, \gamma_{ijk}, \ldots,$$

so entsteht jeder Merkmalwert a_{ijk} (wenn man sich auf drei Stufen beschränkt) dadurch, daß man dem Gesamtmittelwert a die Zufallsabweihungen der drei Stufen überlagert, also

$$a_{ijk} = a + \alpha_i + \beta_{ij} + \gamma_{ijk}. \tag{1}$$

Dabei gilt für i, j und k

$$\begin{aligned} \mathrm{i} &= 1\,;\ 2\,;\ 3\,;\ \ldots \quad \mathrm{N}_1; \\ \mathrm{j} &= 1\,;\ 2\,;\ 3\,;\ \ldots \quad \mathrm{N}_2; \\ \mathrm{k} &= 1\,;\ 2\,;\ 3\,;\ \ldots \quad \mathrm{N}_3. \end{aligned} \tag{2}$$

Für die *Mittelwerte* hat man

$$\begin{aligned} \mathscr{E}\{\alpha_i\} &= 0\,; \\ \mathscr{E}_{i=\mathrm{konst}}\{\beta_{ij}\} &= 0 \qquad \text{für alle } i\,; \\ \mathscr{E}_{(i;j)=\mathrm{konst}}\{\gamma_{ijk}\} &= 0 \qquad \text{für alle } (i\,;j)\,; \end{aligned} \tag{3}$$

und für die *Varianzen* gilt

$$\begin{aligned} \mathscr{V}\{\alpha_i\} &= \sigma_1^2 = V_1 \\ \mathscr{V}_{i=\mathrm{konst}}\{\beta_{ij}\} &= \sigma_2^2 = V_2 \qquad \text{für alle } i\,; \\ \mathscr{V}_{(i;j)=\mathrm{konst}}\{\gamma_{ijk}\} &= \sigma_3^2 = V_3 \qquad \text{für alle } (i\,;j)\,. \end{aligned} \tag{4}$$

Das dreistufige Modell ist demnach durch vier Parameter, einen Mittelwert a und drei Varianzen σ_1^2, σ_2^2 und σ_3^2, gekennzeichnet.

Die vier unbekannten Modellparameter bestimmt man aus einer *Zufallsprobe* der Größe $n_1 n_2 n_3$, bei der man in der ersten Stufe $n_1 < N_1$ Einheiten auswählt. Aus jeder dieser n_1 *Einheiten erster Stufe* entnimmt man $n_2 < N_2$ *Einheiten zweiter Stufe* usw. Insgesamt hat man (bei drei Stufen) $n_1 n_2 n_3$ Einheiten dritter Stufe zur Verfügung. Nur die Merkmalwerte x_{ijk} dieser Einheiten werden gemessen oder beobachtet. Für die Probe gilt

$$\begin{aligned} \mathrm{i} &= 1\,;\ 2\,;\ \ldots\quad \mathrm{n}_1\,; \\ \mathrm{j} &= 1\,;\ 2\,;\ \ldots\quad \mathrm{n}_2\,; \\ \mathrm{k} &= 1\,;\ 2\,;\ \ldots\quad \mathrm{n}_3\,. \end{aligned} \tag{5}$$

Selbstverständlich stimmen die x_{ijk} mit einem Teil der a_{ijk} überein.

Die varianzanalytische Auswertung der Meßreihe liefert Schätzwerte für die drei Varianzen, deren Signifikanz man (in bekannter Weise) mit dem F-Test beurteilt. Für die weiteren Überlegungen sei vorausgesetzt, daß man die Hypothesen $\sigma_1^2 \equiv V_1 = 0$ und $\sigma_2^2 \equiv V_2 = 0$ *verwerfen* muß. Die Varianzen σ_ϱ^2 aller Stufen ϱ sollen also vorhanden sein.

Der Schätzwert $\bar{x}$ für den unbekannten Mittelwert a ist der Mittelwert aller Beobachtungen,

$$\bar{x} = \frac{1}{n_1 n_2 n_3} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk}\,. \tag{6}$$

Die *Genauigkeit* dieses Schätzwerts läßt sich durch den Erwartungswert seiner *Varianz* beurteilen. Dabei hat man *zwei Fälle* zu unterscheiden:

1. Man will mit der Versuchsreihe die *endliche Gesamtheit* $N_1 N_2 N_3$ beurteilen, d. h. etwa eine endliche vorgelegte *Liefermenge*.

2. Man faßt diese Gesamtheit nur als *die zufällige Realisierung* eines hinter ihr stehenden »unendlichen« Erzeugungsvorgangs auf, der beurteilt werden soll; d. h. nicht eine bestimmte Liefermenge, sondern der Fertigungsvorgang, aus dem sie entstanden ist, wird untersucht. Im ersten Falle gilt

$$\mathscr{V}\{\bar{x}\} = \frac{V_1}{n_1}\,\frac{N_1 - n_1}{N_1 - 1} + \frac{V_2}{n_1 n_2}\,\frac{N_2 - n_2}{N_2 - 1} + \frac{V_3}{n_1 n_2 n_3}\,\frac{N_3 - n_3}{N_3 - 1}\,, \tag{7}$$

während im zweiten Falle alle *Endlichkeitsfaktoren* $(N_\varrho - n_\varrho)/(N_\varrho - 1)$ in der letzten Gleichung wegfallen.

In Gleichung (7) wird die Varianz V_ϱ jeder Stufe ϱ im Verhältnis zur Gesamtzahl $\prod_{\lambda=1}^{\varrho} n_\lambda$ der in dieser Stufe ausgewählten Einheiten herabgesetzt. Die Endlichkeitsfaktoren $(N_\varrho - n_\varrho)/(N_\varrho - 1)$ berücksichtigen die Tatsache, daß die endliche Gesamtheit mit wachsender Zahl n_ϱ der ausgewählten Einheiten schließlich ausgeschöpft wird. Es gilt also

$$1 \leq n_\varrho \leq N_\varrho\,. \tag{8}$$

Die Größe der Varianz $\mathfrak{V}\{\bar{x}\}$ hängt wesentlich davon ab, wie man den Gesamtaufwand auf die einzelnen Stufen verteilt. Im folgenden wird untersucht, wie man die Probengröße n_ϱ in der Stufe ϱ wählen muß, damit bei

fest vorgegebenen Gesamtkosten	vorgeschriebener Varianz
$K = K_0$	$\mathfrak{V}\{\bar{x}\} = V_0$ (9)
der Untersuchung die Varianz $\mathfrak{V}\{\bar{x}\}$ des Schätzwerts $\bar{x}$ möglichst klein wird;	des Schätzwerts $\bar{x}$ die Gesamtkosten K der Untersuchung möglichst klein werden;
$\mathfrak{V}\{\bar{x}\} = \text{Min}$	$K = \text{Min}\,.$ (10)

Die Kosten für die Entnahme *einer* Einheit der Stufe ϱ werden mit c_ϱ bezeichnet. Dann sind die *Gesamtkosten*

$$K = c_1 n_1 + c_2 n_1 n_2 + c_3 n_1 n_2 n_3\,. \tag{11}$$

2. BEURTEILUNG DES ERZEUGUNGSVORGANGS (BEURTEILUNG DER FERTIGUNG). NÄHERUNGSLÖSUNG FÜR $n_\varrho \ll N_\varrho$

Wenn die Zahl n_ϱ der in der Stufe ϱ *entnommenen* Einheiten klein gegen die Zahl N_ϱ der dort *vorhandenen* Einheiten bleibt, oder wenn nicht die Liefermenge, sondern der Fertigungsvorgang beurteilt werden soll, so gilt

$$\mathfrak{V}\{\bar{x}\} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1 n_2} + \frac{\sigma_3^2}{n_1 n_2 n_3}\,. \tag{12}$$

Die weitere Rechnung wird wie bisher auf ein dreistufiges Modell beschränkt, da alle Gesetzmäßigkeiten bereits damit gut zum Ausdruck kommen. Mehr Stufen erfordern nicht mehr Denkaufwand, sondern nur mehr Schreibarbeit. Die Auswahlzahlen $(n_1;\ n_2;\ n_3)$ werden vorerst mit $(x;\ y;\ z)$ bezeichnet und als stetige Veränderliche betrachtet. Ferner sei

$$V(x\,;y\,;z) = \frac{\sigma_1^2}{x} + \frac{\sigma_2^2}{xy} + \frac{\sigma_3^2}{xyz} \tag{13}$$

die *Varianzfunktion* und

$$K(x\,;y\,;z) = c_1 x + c_2 xy + c_3 xyz \tag{14}$$

die *Kostenfunktion*.

Die beiden eingangs gestellten Aufgaben lauten damit

$$V(x\,;y\,;z) = \text{Min} \quad | \quad K(x\,;y\,;z) = \text{Min} \tag{15}$$

unter der Nebenbedingung

$$K(x\,;y\,;z) - K_0 = 0 \quad | \quad V(x\,;y\,;z) - V_0 = 0\,. \tag{16}$$

Bezeichnet man mit λ_x^2 und λ_y^2 den jeweiligen Multiplikator (von Lagrange), so findet man die notwendigen Bedingungen für ein Minimum aus der Gleichung

$$V + \lambda_x^2 K = \text{Min} \quad | \quad K + \lambda_y^2 V = \text{Min}. \tag{17}$$

Setzt man

$$\lambda_x^2 \lambda_y^2 = 1 \quad \text{und} \quad \lambda_x^2 = \lambda^2,\ \lambda^2 \neq 0\,, \tag{18}$$

so muß in *beiden* Fällen gelten

$$V_x + \lambda^2 K_x = 0\,; \tag{19}$$

$$V_y + \lambda^2 K_y = 0\,; \tag{20}$$

$$V_z + \lambda^2 K_z = 0\,, \tag{21}$$

wobei V_x, V_y, ... und K_x, K_y, ... die partiellen Ableitungen der Funktionen $V(x; y; z)$ und $K(x; y; z)$ nach x, y, ... bedeuten.

Die drei Gleichungen (19), (20) und (21) liefern der Reihe nach

$$\frac{\sigma_1^2}{x} + \frac{\sigma_2^2}{xy} + \frac{\sigma_3^2}{xyz} = \lambda^2(c_1 x + c_2 xy + c_3 xyz)\,; \tag{22}$$

$$\frac{\sigma_2^2}{xy} + \frac{\sigma_3^2}{xyz} = \lambda^2(c_2 xy + c_3 xyz)\,; \tag{23}$$

$$\frac{\sigma_3^2}{xyz} = \lambda^2(c_3 xyz)\,. \tag{24}$$

Bildet man die *Differenzen*, so findet man

$$\sigma_1^2 = c_1(\lambda x)^2 \quad \text{oder} \quad \lambda x = \frac{\sigma_1}{\sqrt{c_1}}\,; \tag{25}$$

$$\sigma_2^2 = c_2(\lambda xy)^2 \quad \text{oder} \quad \lambda xy = \frac{\sigma_2}{\sqrt{c_2}}\,; \tag{26}$$

$$\sigma_3^2 = c_3(\lambda xyz)^2 \quad \text{oder} \quad \lambda xyz = \frac{\sigma_3}{\sqrt{c_3}}\,. \tag{27}$$

Bildet man schließlich die *Quotienten*, so lassen sich $y = n_2$ und $z = n_3$ sofort bestimmen zu

$$y = n_2 = \sqrt{\frac{c_1}{c_2}}\left(\frac{\sigma_2}{\sigma_1}\right)\,; \tag{28}$$

$$z = n_3 = \sqrt{\frac{c_2}{c_3}}\left(\frac{\sigma_3}{\sigma_2}\right)\,. \tag{29}$$

Das Wertepaar $(y;\ z) = (n_2;\ n_3)$ hängt nur vom *Verhältnis* $(c_{\varrho-1}/c_\varrho)$ *der Kostenparameter* und vom *Verhältnis* $(\sigma_\varrho/\sigma_{\varrho-1})$ *der Standardabweichungen* ab, die für die Stufen 1, 2 und 3 des Modells gelten. n_2 und n_3 sind *unabhängig* von den für die Gesamtkosten K oder für die Varianz $\mathscr{V}\{\bar{x}\}$ vorgeschriebenen Werten K_0 oder V_0.

Zur Berechnung von λ setzt man die Ausdrücke für (λx), (λxy) und (λxyz) aus (25), (26) und (27) in die

Kostenfunktion (16)	Varianzfunktion (16)

ein und findet

$$\sqrt{c_1}\,\sigma_1 + \sqrt{c_2}\,\sigma_2 + \sqrt{c_3}\,\sigma_3 = \lambda K_0 \quad \Big| \quad \lambda(\sqrt{c_1}\,\sigma_1 + \sqrt{c_2}\sigma_2 + \sqrt{c_3}\sigma_3) = V_0 \tag{30}$$

Erklärt man die Hilfsgröße S_{123} durch

$$S_{123} = \sqrt{c_1}\,\sigma_1 + \sqrt{c_2}\,\sigma_2 + \sqrt{c_3}\,\sigma_3 \tag{31}$$

dann wird λ zu

$$\lambda = \frac{S_{123}}{K_0}; \quad \Big| \quad \lambda = \frac{V_0}{S_{123}}. \tag{32}$$

Damit läßt sich nunmehr aus (25) auch $x = n_1$ berechnen. Man findet

$$x = n_1 = \frac{\sigma_1}{\sqrt{c_1}} \frac{K_0}{S_{123}} \quad \Big| \quad x = n_1 = \frac{\sigma_1}{\sqrt{c_1}} \frac{S_{123}}{V_0}. \tag{33}$$

Zur Ermittlung der kleinsten

Varianz bei vorgegebenen Gesamtkosten $K = K_0$	Gesamtkosten bei vorgeschriebener Varianz $\mathscr{V}\{\bar{x}\} = V_0$

berechnet man aus (26) und (27) zweckmäßig auch noch die Produkte xy und xyz. Es wird

$$xy = n_1 n_2 = \frac{\sigma_2}{\sqrt{c_2}} \frac{K_0}{S_{123}} \quad \Big| \quad xy = n_1 n_2 = \frac{\sigma_2}{\sqrt{c_2}} \frac{S_{123}}{V_0}$$

$$xyz = n_1 n_2 n_3 = \frac{\sigma_3}{\sqrt{c_3}} \frac{K_0}{S_{123}} \quad \Big| \quad xyz = n_1 n_2 n_3 = \frac{\sigma_3}{\sqrt{c_3}} \frac{S_{123}}{V_0}.$$

Setzt man diese Ausdrücke in die

Varianzfunktion (13)	Kostenfunktion (14)

ein, so findet man schließlich die erreichbaren Minimalwerte zu

$$\mathscr{V}_{\min}\{\bar{x}\} = \frac{S_{123}^2}{K_0} \quad \Big| \quad K_{\min} = \frac{S_{123}^2}{V_0}. \tag{34}$$

Je größer der verfügbare Gesamtbetrag K_0 ist, umso genauer läßt sich $\bar{x}$ aus der Untersuchung bestimmen.

Je genauer man den Mittelwert $\bar{x}$ erfassen will, umso größer ist der aufzuwendende Gesamtbetrag $K_{\min}$ für die Untersuchung.

In beiden Fällen hat das Produkt aus dem vorgegebenen Wert und dem erreichbaren Wert den gleichen Betrag

$$\mathfrak{V}_{\min} \cdot K_0 = K_{\min} \cdot V_0 = S_{123}^2 = \text{konst}. \tag{35}$$

Das kommt bereits in Gleichung (22) zum Ausdruck, denn man kann ihr die Form geben

$$\mathfrak{V}\{\bar{x}\} = \lambda^2 K. \tag{36}$$

Setzt man hier λ aus (32) ein, so kommt man auf die Gleichungen (34) zurück.

Ersetzt man die Standardabweichungen σ_ϱ durch $\alpha\sigma_\varrho$ und die Kostenparameter c_ϱ durch $\beta^2 c_\varrho$, wobei die Faktoren α und β^2 fest gewählte Werte darstellen, so geht

$$\begin{aligned} V \quad &\text{über in} \quad V' &&= \alpha^2 V, \\ K \quad &\text{in} \quad K' &&= \beta^2 K \\ \text{und} \quad S_{123} \quad &\text{in} \quad S'_{123} &&= \alpha\beta S_{123}. \end{aligned}$$

Aus den Gleichungen (28), (29) und (33) entnimmt man, daß die Lösung dabei völlig ungeändert bleibt; es gilt $n'_\varrho = n_\varrho$.

Die wirtschaftlich günstigsten Verhältnisse erreicht man, wenn man den Aufwand n_2 und n_3 für die Stufen 2 und 3 aus den Gleichungen (28) und (29) berechnet. Zu dem Zweck muß man Schätzwerte für die Varianzen σ_ϱ^2 aus einem *Vorversuch* bestimmen, bei dem man die $n_\varrho \geq 2$ an sich beliebig wählen darf. Damit jedoch bei der Varianzanalyse die den drei Stufen zugeordneten Freiheitsgrade wenigstens »von gleicher Größenordnung« werden, setzt man im Vorversuch zweckmäßig $n_2 = 2$ und $n_3 = 2$. Dann werden die den Stufen 1, 2 und 3 zugeordneten Zahlen der Freiheitsgrade

$$f_{\mathrm{I}} = n_1 - 1 \; ; \; f_{\mathrm{II}} = n_1(n_2 - 1) = n_1 \; ; \; f_{\mathrm{III}} = n_1 n_2 (n_3 - 1) = 2n_1 .$$

Mißt man die Genauigkeit G des Schätzwerts $\bar{x}$ durch den Kehrwert $1/\sigma\{\bar{x}\}$ der Standardabweichung $\sigma\{\bar{x}\}$, so folgt aus (34)

$$G = \frac{1}{S_{123}} \sqrt{K_0}. \tag{37}$$

Die Genauigkeit G wächst also (leider) nur mit der Wurzel aus dem Gesamtaufwand K_0.

Während $y = n_2$ und $z = n_3$ feste Werte sind, die durch die Modellparameter bestimmt werden, wächst $x = n_1$ linear mit K_0 an oder sinkt mit dem Kehrwert $1/V_0$. In beiden Fällen kann $x = n_1$ (rechnerisch) schließlich über alle Grenzen wachsen. Die im vorstehenden hergeleitete Lösung bleibt

jedoch nur sinnvoll, wenn alle $n_\varrho \ll N_\varrho$ sind. Wird diese Bedingung verletzt, was in der ersten Stufe bei n_1 mit wachsendem K_0 oder abnehmendem V_0 immer möglich ist, so muß man auf die Lösung des Abschnitts 4 zurückgreifen.

3. EIN SONDERFALL

Ist für das Modell insbesondere

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \quad \text{und} \quad c_1 \leq c_2 \leq c_3 \tag{38}$$

was in den Anwendungen durchaus vorkommt, dann folgt aus (28) und (29) sofort

$$y \leq 1 \quad \text{und} \quad z \leq 1 .$$

Verwirklichen kann man jedoch nur

$$n_2 = 1 \quad \text{und} \quad n_3 = 1 .$$

In der vorausgehenden Rechnung wurden die Auswahlzahlen x, y und z als *stetige* Veränderliche betrachtet. In Wirklichkeit ändern sie sich natürlich *sprunghaft* von 1 zu 1. Man kann also nicht immer die günstige Lösung $V_{\min}$ für die Varianz bei gegebenem K_0 (oder $K_{\min}$ für die Kosten bei gegebenem V_0) erreichen, sondern dieser Lösung nur mehr oder weniger nahe kommen. Sind $(x; y; z)$ alle größer als 1, so wählt man für $(n_1; n_2; n_3)$ die zu $(x; y; z)$ benachbarten ganzen Zahlen. Ist jedoch (auf Grund der Modellparameter σ_ϱ und der Kostenfaktoren c_ϱ) y und z kleiner als 1, so muß man die *»rechnerische« Kostenfunktion*

$$K = c_1 x + c_2 xy + c_3 xyz$$

durch die *»wirkliche« Kostenfunktion*

$$K = (c_1 + c_2 + c_3)x \tag{39}$$

ersetzen, da man in den Stufen 2 und 3 mindestens *eine* Einheit auswählen muß. Anstelle von (33) tritt dann

$$x = \frac{K_0}{c_1 + c_2 + c_3} = n_1 . \tag{40}$$

Rechnerisch ist für $(x; y; z)$ nach (34)

$$\mathscr{V}_{\min}\{\bar{x}\} = \frac{S_{123}^2}{K_0} = \frac{1}{K_0}(\sqrt{c_1}\sigma_1 + \sqrt{c_2}\sigma_2 + \sqrt{c_3}\sigma_3)^2 . \tag{41}$$

Tatsächlich kann man mit $(n_1; 1; 1)$ aber nur

$$\mathscr{V}_{\text{err}}\{\bar{x}\} = \frac{1}{n_1}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) = \frac{1}{K_0}(c_1 + c_2 + c_3)(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) \tag{42}$$

erreichen. Für den Sonderfall $\sigma_\varrho = \text{konst} = \sigma$ und $c_\varrho = \text{konst} = c$ gilt

$$\mathscr{V}_{\min}\{\bar{x}\} = \frac{9c\sigma^2}{K_0} = \mathscr{V}_{\text{err}}\{\bar{x}\}. \tag{43}$$

Da beide Ausdrücke übereinstimmen, ist das Minimum erreichbar. Im allgemeinen Falle kann jedoch der Unterschied $\Delta = (\mathscr{V}_{\text{err}} - \mathscr{V}_{\min})$ oder

$$\Delta = \frac{1}{K_0}[(\sqrt{c_1}\sigma_2 - \sqrt{c_2}\sigma_1)^2 + (\sqrt{c_2}\sigma_3 - \sqrt{c_3}\sigma_2)^2 + (\sqrt{c_3}\sigma_1 - \sqrt{c_1}\sigma_3)^2] \tag{44}$$

beträchtlich werden, wenn das Modell durch $\sigma_1 > \sigma_2 > \sigma_3$ und gleichzeitig $c_1 < c_2 < c_3$ gekennzeichnet ist.

Hat man beispielsweise

$$\sigma_1 = 4\,;\ \sigma_2 = 2\,;\ \sigma_3 = 1 \quad \text{und} \quad c_1 = 1\,;\ c_2 = 4\,;\ c_3 = 16\,,$$

so wird $y = \frac{1}{4}$ und $z = \frac{1}{4}$. Man muß also $n_2 = 1$ und $n_3 = 1$ wählen

und hat als wirkliche Kostenfunktion (39)

$$K = 21x\,.$$

Es wird bei gegebenem K_0

$$\mathscr{V}_{\min}\{\bar{x}\} = \frac{S_{123}^2}{K_0} = \frac{144}{K_0} \quad \text{und} \quad \mathscr{V}_{\text{err}}\{\bar{x}\} = \frac{441}{K_0}\,,$$

also

$$\mathscr{V}_{\text{err}} \approx 3\mathscr{V}_{\min}\,.$$

Ist dagegen »umgekehrt«

$$\sigma_1 = 1\,;\ \sigma_2 = 2\,;\ \sigma_3 = 4 \quad \text{und} \quad c_1 = 16\,;\ c_2 = 4\,;\ c_3 = 1\,,$$

so wird $y = 4 = n_2$ und $z = 4 = n_3$. Jetzt darf man die »rechnerische« Kostenfunktion verwenden und findet mit $S_{123} = 12$

$$K = 48x \quad \text{und} \quad \mathscr{V}_{\min}\{\bar{x}\} = \mathscr{V}_{\text{err}}\{\bar{x}\} = \frac{144}{K_0}\,.$$

Für $K_0 = 144$ ist $x = 3 = n_1$ und $\mathscr{V}_{\min}\{\bar{x}\} = 1$.

Die Gesamtkosten $K_0 = 144$ sind durch die in der folgenden Zahlentafel zusammengestellten Wertetripel $(n_1;\ n_2;\ n_3)$ zu verwirklichen (Tabelle I). Ersichtlich ist bei allen diesen Lösungen die Varianz größer als der günstigste Wert $\mathscr{V}_{\min}\{\bar{x}\} = 1$, der für das Wertetripel (3; 4; 4) zustande kommt.

TABELLE I

$c_1 = 16;\ c_2 = 4;\ c_3 = 1$			
$\sigma_1^2 = 1;\ \sigma_2^2 = 4;\ \sigma_3^2 = 16$			
n_1	n_2	n_3	$V\{\bar{x}\}$
1	1	124	5,13
	2	60	3,13
	4	28	2,14
	8	12	1,67
	16	4	1,50
2	1	52	2,65
	2	24	1,67
	4	10	1,20
	7	4	1,07
	8	3	1,08
3	1	28	1,86
	2	12	1,22
	4	4	1,00
4	1	16	1,50
	2	6	1,08
	4	1	1,50
6	1	4	1,50

4. BEURTEILUNG DER LIEFERMENGE. LÖSUNG FÜR BELIEBIGE $n_\varrho \leq N_\varrho$

In diesem Falle geht man von Gleichung (7) aus. Ersetzt man $(n_1; n_2; n_3)$ zunächst wieder durch $(x; y; z)$, so lautet die Varianzfunktion jetzt

$$V(x;y;z) = \frac{\sigma_1^2}{N_1 - 1}\left(\frac{N_1}{x} - 1\right) + \frac{1}{x}\,\frac{\sigma_2^2}{N_2 - 1}\left(\frac{N_2}{y} - 1\right) + \\ + \frac{1}{xy}\,\frac{\sigma_3^2}{N_3 - 1}\left(\frac{N_3}{z} - 1\right). \tag{45}$$

Die Kostenfunktion $K(x; y; z)$ bleibt gegen (11) ungeändert. Jedoch sei vorausgesetzt, daß der verfügbare Gesamtbetrag

$$K_0 \leq c_1 N_1 + c_2 N_1 N_2 + c_3 N_1 N_2 N_3 = K_N \tag{46}$$

bleibt. Ist im Grenzfalle $K_0 = K_N$, so wählt man $x = N_1$; $y = N_2$; $z = N_3$ und findet aus (7) $\mathfrak{V}\{\bar{x}\} = 0$. Man wäre für $K_0 = K_N$ in der Lage, in allen Stufen alle Einheiten zu prüfen (Vollprüfung) und würde damit den rich-

tigen Mittelwert $\bar{x} = a$ finden. Für $K_0 > K_N$ bliebe sogar noch Geld übrig. In allen praktischen Fällen ist K_0 *wesentlich kleiner* als K_N. Die Gleichungen (19) bis (21) liefern jetzt der Reihe nach

$$\frac{1}{x}\frac{\sigma_1^2}{(N_1-1)}N_1+\frac{1}{x}\frac{\sigma_2^2}{(N_2-1)}\left(\frac{N_2}{y}-1\right)+\frac{1}{xy}\frac{\sigma_3^2}{(N_3-1)}\left(\frac{N_3}{z}-1\right)= \\ =\lambda^2(c_1x+c_2xy+c_3xyz)\,; \tag{47}$$

$$\frac{1}{xy}\frac{\sigma_2^2}{(N_2-1)}N_2+\frac{1}{xy}\frac{\sigma_3^2}{(N_3-1)}\left(\frac{N_3}{z}-1\right)=\lambda^2(c_2xy+c_3xyz)\,; \tag{48}$$

$$\frac{1}{xyz}\frac{\sigma_3^2}{(N_3-1)}N_3=\lambda^2(c_3xyz)\,. \tag{49}$$

Für später ist es zweckmäßig, Gleichung (47) in der Form

$$\mathscr{V}\{\bar{x}\}+\frac{\sigma_1^2}{N_1-1}=\lambda^2K \tag{50}$$

zu schreiben.

Bildet man die *Differenzen* (47) — (48) und (48) — (49), so findet man

$$\frac{N_1}{N_1-1}\sigma_1^2-\frac{\sigma_2^2}{N_2-1}=c_1(\lambda x)^2\,; \tag{51}$$

$$\frac{N_2}{N_2-1}\sigma_2^2-\frac{\sigma_3^2}{N_3-1}=c_2(\lambda xy)^2\,; \tag{52}$$

$$\frac{N_3}{N_3-1}\sigma_3^2=c_3(\lambda xyz)^2\,. \tag{53}$$

Erklärt man die Hilfsgrößen W_ϱ^2 durch die Gleichungen

$$W_1^2=\frac{N_1}{N_1-1}-\frac{(\sigma_2/\sigma_1)^2}{N_2-1}\approx 1+\frac{1}{N_1}-\frac{(\sigma_2/\sigma_1)^2}{N_2}\,; \tag{54}$$

$$W_2^2=\frac{N_2}{N_2-1}-\frac{(\sigma_3/\sigma_2)^2}{N_3-1}\approx 1+\frac{1}{N_2}-\frac{(\sigma_3/\sigma_2)^2}{N_3}\,; \tag{55}$$

$$W_3^2=\frac{N_3}{N_3-1}\approx 1+\frac{1}{N_3}\,, \tag{56}$$

wobei

$$\lim_{N_\varrho ; N_{\varrho+1} \to \infty} W_\varrho^2 = 1 \tag{57}$$

ist, so gilt

$$\sigma_1 W_1 = \sqrt{c_1}\,\lambda x\,, \tag{58}$$

$$\sigma_2 W_2 = \sqrt{c_2}\,(\lambda x y)\,, \tag{59}$$

$$\sigma_3 W_3 = \sqrt{c_3}\,(\lambda x y z)\,. \tag{60}$$

Auch die Varianzfunktion (45) läßt sich damit umgestalten zu

$$V(x; y; z) + \frac{\sigma_1^2}{N_1 - 1} = \frac{\sigma_1^2}{x} W_1^2 + \frac{\sigma_2^2}{xy} W_2^2 + \frac{\sigma_3^2}{xyz} W_3^2\,. \tag{61}$$

Bildet man die *Quotienten* (59)/(58) und (60)/(59), so lassen sich $y = n_2$ und $z = n_3$ sofort berechnen zu

$$y = n_2 = \sqrt{\frac{c_1}{c_2}} \left(\frac{\sigma_2}{\sigma_1}\right) \left(\frac{W_2}{W_1}\right) = y_\infty \frac{W_2}{W_1}\,, \tag{62}$$

$$z = n_3 = \sqrt{\frac{c_2}{c_3}} \left(\frac{\sigma_3}{\sigma_2}\right) \left(\frac{W_3}{W_2}\right) = z_\infty \frac{W_3}{W_2}\,. \tag{63}$$

Man muß demnach die Näherungswerte aus (28) und (29), die hier mit y_∞ und z_∞ bezeichnet wurden, mit den Verhältnissen $W_\varrho / W_{\varrho-1}$ multiplizieren. Da die Hilfsgrößen W_ϱ meist nahe bei 1 liegen, sind die Unterschiede zwischen den Näherungswerten $(y_\infty; z_\infty)$ und den genauen Werten $(y; z)$ im allgemeinen nicht groß und praktisch ohne Bedeutung. Auch jetzt hängt das Wertepaar $(n_2 ; n_3)$ *nicht* von den vorgegebenen Werten K_0 oder V_0 ab, sondern nur von den Verhältnissen $(c_{\varrho-1}/c_\varrho)$, $(\sigma_\varrho/\sigma_{\varrho-1})$ und $(\sigma_{\varrho+1}/\sigma_\varrho)$.

Zur Berechnung von λ setzt man die Ausdrücke für (λx), (λxy) und (λxyz) aus (58) bis (60) in die

Kostenfunktion (14)	Varianzfunktion (61)

ein und findet

$$\lambda = \frac{T_{123}}{K_0} \quad \Big| \quad \lambda = \frac{V_0 + [\sigma_1^2/(N_1 - 1)]}{T_{123}}\,, \tag{64}$$

wobei die Hilfsgröße T_{123} durch

$$T_{123} = \sqrt{c_1}\,\sigma_1 W_1 + \sqrt{c_2}\,\sigma_2 W_2 + \sqrt{c_3}\,\sigma_3 W_3 \tag{65}$$

erklärt ist. Aus (25) folgt schließlich

$$x = n_1 = \frac{\sigma_1}{\sqrt{c_1}} W_1 \frac{K_0}{T_{123}} \quad \Big| \quad x = n_1 = \frac{\sigma_1}{\sqrt{c_1}} W_1 \frac{T_{123}}{V_0 + [\sigma_1^2/(N_1 - 1)]}\,. \tag{66}$$

Aus (50) erhält man mit dem jeweils entsprechenden λ^2 aus (64) die

kleinste Varianz $\mathfrak{V}_{\min}$ bei vorgegebenen Gesamtkosten $K = K_0$	kleinsten Gesamtkosten $K_{\min}$ bei vorgeschriebener Varianz $\mathfrak{V}\{\bar{x}\} = V_0$,
$\mathfrak{V}_{\min}\{\bar{x}\} = \dfrac{T_{123}^2}{K_0} - \dfrac{\sigma_1^2}{N_1 - 1}$	$K_{\min} = \dfrac{T_{123}^2}{V_0 + [\sigma_1^2/(N_1 - 1)]}$. (67)
Das Verhalten von	
$\mathfrak{V}_{\min}\{\bar{x}\}$ für $K_0 \to K_N$	$K_{\min}$ für $V_0 \to 0$

wird später noch untersucht.

Ersetzt man die Standardabweichungen σ_ϱ durch $\sigma'_\varrho = \alpha\sigma_\varrho$ mit $\alpha = $ = konst. und die Kostenfaktoren c_ϱ durch $c'_\varrho = \beta^2 c_\varrho$ mit $\beta^2 =$ konst. so ändern sich die Hilfsgrößen W_ϱ nach (54) bis (56) nicht. Aus K wird $K' = \beta^2 K$ und aus T_{123} wird $T'_{123} = \alpha\beta T_{123}$. Auch jetzt bleibt demnach $n'_\varrho = n_\varrho$ ungeändert, wie man unmittelbar aus (62), (63) und (66) erkennt.

5. DIE GRENZÜBERGÄNGE $K_0 \to K_N$ UND $V_0 \to 0$

Die bisher mit $K(x; y; z)$ und $V(x; y; z)$ bezeichneten Funktionen für die Kosten und die Varianz und die fest vorgegebenen Sollwerte K_0 und V_0 werden im folgenden mit $K_1(x; y; z)$, $V_1(x; y; z)$, K_{10} und V_{10} bezeichnet. Der zusätzliche Index 1 weist auf die *Stufe* hin, die mit wachsenden Kosten K_{10} oder abnehmender Varianz V_{10} *abgebaut* wird.

Läßt man	
$K_0 \equiv K_{10}$ unbegrenzt wachsen,	$V_0 \equiv V_{10}$ unbegrenzt abnehmen,

so wächst $\varkappa = n_1$ nach (66) monoton an. Das ist möglich, bis n_1 den Grenzwert N_1 erreicht. Für $n_1 = N_1$ verschwindet der Varianzanteil der Stufe 1, denn

$$\frac{\sigma_1^2}{n_1}\,\frac{N_1 - n_1}{N_1 - 1} = 0 \qquad \text{für } n_1 = N_1 . \tag{68}$$

Dann ist

$K_{10} = K_{10}^* = \dfrac{\sqrt{c_1}N_1}{\sigma_1 W_1}\, T_{123}$	$V_{10} = V_{10}^* = \dfrac{\sigma_1 W_1}{\sqrt{c_1}N_1}\, T_{123} - \dfrac{\sigma_1^2}{N_1 - 1}$. (69)

Läßt man jetzt K_{10} noch weiter wachsen oder V_{10} noch weiter abnehmen, so bleibt n_1 fest.

$$n_1 = \text{konst} = N_1 , \tag{70}$$

und das Minimalproblem geht in die entsprechende Aufgabe für ein Modell mit nur zwei Stufen über. Es übernimmt y jetzt die Rolle von x und z die

von y, wobei die Varianz der Stufe 2 abgebaut wird. Für die *neue Varianzfunktion* $V_2(y; z)$ erhält man aus (45) mit $x = N_1$

$$N_1 V_1(N_1; y; z) = V_2(y; z) = \frac{\sigma_2^2}{N_2 - 1}\left(\frac{N_2}{y} - 1\right) + \frac{1}{y}\frac{\sigma_3^2}{N_3 - 1}\left(\frac{N_3}{z} - 1\right). \quad (71)$$

Die neue Kostenfunktion $K_2(y; z)$ folgt aus

$$K_1(N_1; y; z) = c_1 N_1 + N_1(c_2 y + c_3 y z) \quad (72)$$

zu

$$K_2(y; z) = c_2 y + c_3 y z = \frac{K_1(N_1; y; z)}{N_1} - c_1 . \quad (73)$$

Die Lösung dieses Zwei-Stufen-Problems findet man aus den vorausgehenden Gleichungen, wenn man den Index 1 durch 2, 2 durch 3, ferner x durch y und y durch z ersetzt. Man findet aus (66) und (62)

$$y = n_2 = \frac{\sigma_2}{\sqrt{c_2}} W_2 \frac{K_{20}}{T_{23}} \quad \Big| \quad y = n_2 = \frac{\sigma_2}{\sqrt{c_2}} W_2 \frac{T_{23}}{V_{20} + [\sigma_2^2/(N_2 - 1)]} \quad (74)$$

$$z = n_3 = \sqrt{\frac{c_2}{c_3}} \frac{\sigma_3}{\sigma_2} \frac{W_3}{W_2}, \quad (75)$$

d. h. $z = n_3$ bleibt gegenüber der dreistufigen Lösung zunächst ungeändert. Läßt man jetzt

K_{20} weiter wachsen, | V_{20} weiter abnehmen,

so wächst y monoton an, bis es den Grenzwert $y = n_2 = N_2$ erreicht. Dann ist nach (74)

$$K_{20} = K_{20}^* = \frac{\sqrt{c_2} N_2}{\sigma_2 W_2} T_{23} \quad \Big| \quad V_{20} = V_{20}^* = \frac{\sigma_2 W_2}{\sqrt{c_2} N_2} T_{23} - \frac{\sigma_2^2}{(N_2 - 1)}. \quad (76)$$

Jetzt verschwindet auch noch der Varianzanteil der zweiten Stufe,

$$\frac{\sigma_2^2}{n_2}\frac{N_2 - n_2}{N_2 - 1} = 0 \text{ für } y = n_2 = N_2 . \quad (77)$$

Es bleibt allein übrig

$$N_2 V_2(N_2; z) = V_3(z) = \frac{\sigma_3^2}{N_3 - 1}\left(\frac{N_3}{z} - 1\right) \quad (78)$$

und

$$K_2(N_2; z) = c_2 N_2 + N_2 c_3 z$$

oder

$$K_3(z) = c_3 z = \frac{K_2(N_2; z)}{N_2} - c_2 . \quad (79)$$

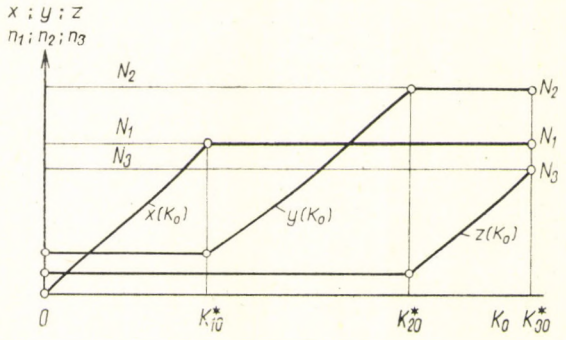


Abb. 1

Bei vorgegebenem Wert $K_3 = K_{30}$ bzw. $V_3 = V_{30}$ wird z

$$z = n_3 = \frac{K_{30}}{c_3} \qquad \left| \qquad z = n_3 = \frac{N_3\sigma_3^2/(N_3-1)}{V_{30} + [\sigma_3^2/(N_3-1)]} \right. \tag{80}$$

oder nach (66) mit 3 als Index anstelle von 1

$$z = \frac{\sigma_3}{\sqrt{c_3}} W_3 \frac{K_{30}}{T_3} \qquad \left| \qquad z = \frac{\sigma_3}{\sqrt{c_3}} W_3 \frac{T_3}{V_{30} + [\sigma_3^2/(N_3-1)]} \right. . \tag{81}$$

K_{30} darf anwachsen bis zum Grenzwert

$$K_{30} = K_{30}^* = c_3 N_3 .$$

V_{30} darf abnehmen bis zum Grenzwert

$$V_{30} = V_{30}^* = 0 . \tag{82}$$

Dann wird $z = n_3 = N_3$.

Damit verschwindet auch noch der Varianzanteil der letzten Stufe,

$$\frac{\sigma_3^2}{n_3} \frac{N_3 - n_3}{N_3 - 1} = 0 \text{ für } z = n_3 = N_3 . \tag{83}$$

Wenn der Grenzzustand $z = n_3 = N_3$ erreicht ist, so berechnet man die Gesamtkosten der Reihe nach aus (79) und (73). Man findet

$$K_3(N_3) = c_3 N_3 ,$$

$$K_2(N_2 ; N_3) = c_2 N_2 + c_3 N_2 N_3 ,$$

$$K_1(N_1 ; N_2 ; N_3) = c_1 N_1 + c_2 N_1 N_2 + c_3 N_1 N_2 N_3 = K_N ,$$

wie es selbstverständlich sein muß.

Der beschriebene Grenzübergang muß (mindestens in der *ersten* Stufe) durchgeführt werden, wenn die Zahl N_1 der Einheiten in dieser Stufe nur »klein« ist. Das Verhalten der Lösung beim Grenzübergang ist aus Abb. 1 ersichtlich.

DIE BERECHNUNG UND PRAKTISCHE ANWENDUNG VON DOPPEL- UND MEHRFACHSTICHPROBENPLÄNEN FÜR VARIABLENPRÜFUNG

von

H. TRAMPEL

VVB Wälzlager und Normteile
Karl-Marx-Stadt, Deutsche Demokratische Republik

Die Produktivität des modernen Industriebetriebes kann nicht allein durch eine stetig weitergreifende Automatisierung der Produktion gesteigert werden, sondern allgemein ausgedrückt nur durch die komplexe Erforschung der im Produktionsprozeß wirkenden technischen und ökonomischen Gesetzmäßigkeiten. Dabei kommt dem Studium der mathematisch-statistischen Gesetzmäßigkeiten, als einer Erscheinungsform technischer und ökonomischer Gesetze, insofern wachsende Bedeutung zu, als die Erkenntnis, daß Qualität niemals erprüft, sondern nur erarbeitet werden kann, mehr und mehr zum Allgemeingut aller am Produktionsprozeß beteiligten wird.

Qualität erarbeiten ist gleichzusetzen mit der Forderung, die Qualität zu steuern. Dieses Vorhaben läßt sich um so wirkungsvoller realisieren, je genauer und je unmittelbarer die zur Steuerung der Produktion benötigten Informationen gewonnen werden. Darin besteht bei den heute vorrangig in Anwendung befindlichen Methoden der mathematischen Statistik eine erhebliche Diskrepanz. Zur Beurteilung der qualitativen Leistungsfähigkeit beliebiger Maschinen sind Methoden der Variablenprüfung notwendig, um zu einer hinreichend genauen Aussage zu gelangen. Diese Verfahren laufen sämtlich darauf hinaus, die Maschinenfunktionsgenauigkeit und dabei insbesondere die sogenannte 'natürliche Toleranz' 6σ zu ermitteln und diese Größe der vorgegebenen Toleranz T gegenüberzustellen.

Der Quotient

$$N_F = \frac{6\sigma}{T}$$

den wir als Niveaus der Fertigung bezeichnen, dient zur Entscheidung über den weiteren Verbleib des betreffenden Produktionsaggregates im Fertigungsprozeß, über die Notwendigkeit und den Zeitpunkt von Generalreparaturen (bei periodischen Untersuchungen ergibt sich, daß N_F gegen 1 konvergiert. Aus der Extrapolation dieser Funktion, die zumeist linear ist, kann der Zeitpunkt der notwendigen Reparatur bestimmt werden.) sowie zur Zuordnung der zu produzierenden Erzeugnisse zu den Maschinen einer bestimmten Abteilung. Stehen zur Durchführung einer Arbeitsoperation mehrere Maschinen zur Verfügung, so erfolgt die Zuordnung direkt proportional, d. h. Erzeugnisse mit sehr kleinen vorgegebenen Toleranzen werden auf Maschinen aufgeteilt mit geringem natürlichen Toleranzbereich 6σ.

Die Bestimmung der Maschinenfunktionsgenauigkeit erfolgt in Abhängigkeit von der gewünschten Genauigkeit der Aussage, der Zahl und der Qualifikation des die Analyse durchführenden Personals. Sie reicht demzufolge von der einfachen Strichliste bis zur Varianz- und Faktorenanalyse.

Unabhängig davon, wie im speziellen Fall die Bestimmung der Maschinenfunktionsgenauigkeit als vielleicht wichtigstes Anwendungsgebiet der Mathematischen Statistik im Industriebetrieb durchgeführt wird, kann sie infolge des relativ hohen Aufwandes nur in Verbindung mit speziellen elektronischen Rechnern, sogenannten Prozeßrechnern, unmittelbar auch zur Beurteilung der Qualität der Erzeugnisse im laufenden Produktionsprozeß verwendet werden.

Da derartige Rechner nur in Ausnahmefällen zur Verfügung stehen, erfolgt die Beurteilung der Qualität zumeist mit Attributstichprobenplänen. Sie besitzen gegenüber entsprechenden Variablenstichprobenplänen, von denen die bekanntesten im MST 414 zusammengefaßt sind, den Vorteil, daß sie keine besonderen Forderungen an die Qualifikation des Kontrollpersonals stellen, diesen Vorteil jedoch durch ein mehrfaches an Kontrollaufwand in der gegenteiligen Richtung ausgleichen.

Die damit vorhandenen Zweispurigkeit zwischen Beurteilung der Maschinenfunktionsgenauigkeit und der Beurteilung der Qualität der laufend auf diesen Maschinen gefertigten Produkte bedeutet erheblichen Mehraufwand.

Dieser Mehraufwand kann nahezu vollständig beseitigt werden durch die Anwendung von Variablenstichprobenplänen, da die Ergebnisse der Variablenprüfung unmittelbar zur Beurteilung und Kontrolle der Maschinenfunktionsgenauigkeit Verwendung finden können.

Die Frage, aus welchem Grunde Variablenstichprobenpläne in so geringem Maße im Einsatz sind, bei derartig offenkundigen Vorteilen, ist nur damit zu beantworten, daß die bestehenden Variablenpläne so hohe Anforderungen hinsichtlich der Qualifikation der Kontrollkräfte und des Umfanges der durchzuführenden Rechenoperationen stellen, daß sie damit in der Praxis scheitern.

An Variablenstichprobenpläne, die die Verbindung der Bestimmung der Maschinenfunktionsgenauigkeit und der Beurteilung der Qualität gestatten, müssen folgende Forderungen gestellt werden:

1. Sie müssen einfach im Aufbau und in der Handhabung sein;
2. Die Variablenstichprobensysteme müssen so aufgebaut sein, daß der Austausch verschiedener Systeme keine Auswirkungen hinsichtlich der Qualitätsbeurteilung hat;
3. Wenn das tatsächliche Qualitätsniveau einer Fertigung über längere Zeit konstant unter dem geforderten Qualitätsniveau liegt, so muß die Beurteilung der Qualität mit reduziertem Aufwand möglich sein;
4. Diese reduzierte Prüfung darf nicht dazu führen, daß die Beurteilung im Bereich des p_β, der sogenannten 'Schlechtgrenze' schwächer wird.

Die theoretische Grundlage der Variablenstichprobenpläne bildet die Normalverteilung, wenn auch zum Teil als hinreichend genaue Verallgemeinerung und Approximation einer Vielzahl spezieller Verteilungsformen.

Die Normalverteilung ist durch die Parameter μ und σ eindeutig bestimmt.

Bei der Prüfung von Stichproben treten an die Stelle der Parameter μ und σ der Grundgesamtheit die entsprechenden Parameter der Stichproben. Nach der Auswahl dieser Parameter unterscheiden sich die einzelnen Variablenstichprobensysteme. Wir kennen demnach:

$\bar{x}$-σ-Pläne
$\bar{x}$-s-Pläne
$\bar{x}$-R-Pläne

Sie sind im MST 414 als D-, B- und C-Pläne gekennzeichnet. Darüber hinaus haben wir neu die

$\tilde{x}$-σ-Pläne und
$\tilde{x}$-QR_i-Pläne

berechnet, die als E- bzw. F-Pläne in eine in der DDR in Vorbereitung befindliche TGL aufgenommen werden.

Die Variablenstichprobenpläne kennen keine Unterscheidung hinsichtlich der Parameter der OC-Kurve, wie sie bei den Attributstichprobenplänen üblich sind.

Der Vorteil bei der Berechnung der Pläne des MST 414 als auch der von uns neu berechneten Pläne besteht darin, daß ihre Operationscharakteristik mit der des MST 105 A (in der DDR überarbeitet als TGL 14 450) identisch ist. Diese Tatsache ermöglicht im Betrieb die wahlweise Verwendung von Attribut- und Variablenstichprobenplänen. Die Variablenstichprobenpläne des gleichen Systems unterscheiden sich lediglich nochmals hinsichtlich ihrer Aussage in Pläne der Form 1 und Pläne der Form 2. Bei den Prüfplänen der Form 1 fällt die Annahmeentscheidung durch Vergleich einer Prüfgröße z (die ich noch näher erläutere) mit der vorgegebenen Toleranz T.

Bei den Prüfplänen der Form 2 fällt die Annahmeentscheidung durch Vergleich des aus dem Qualitätsindex Q bestimmten Fehlerprozentsatzes $p_{\max}$ mit dem im Stichprobenplan angegebenen höchstzulässigen Fehlerprozentsatz M.

Die Arbeitsweise der Variablenstichprobenpläne kann schematisch in folgenden Bildern dargestellt werden.

Die Lage der Toleranzgrenze in Abb. 1 (es wird hier aus Gründen der einfacheren Darstellung nur der Fall einer vorgegebenen oberen Toleranzgrenze erläutert) wird durch den vorgegebenen Fehlerprozentsatz bestimmt. Er ist identisch mit dem Flächenstück rechts der Toleranzgrenze T. Annahme erfolgt, wenn die Prüfgröße $z = \mu + \lambda_T \cdot \sigma$ (im Fall der Grundgesamtheit) kleiner oder gleich der Toleranzgrenze T ist ($z \leq T$).

Die Prüfgröße z, die in dieser Form den Prüfplänen der Form 1 zugrundeliegt, kann in folgender Weise umgeformt werden:

$$\lambda_T = \frac{T - \mu}{\sigma}$$

im Falle der Grundgesamtheit entspricht dem Abszissenwert λ_T die Toleranz T.

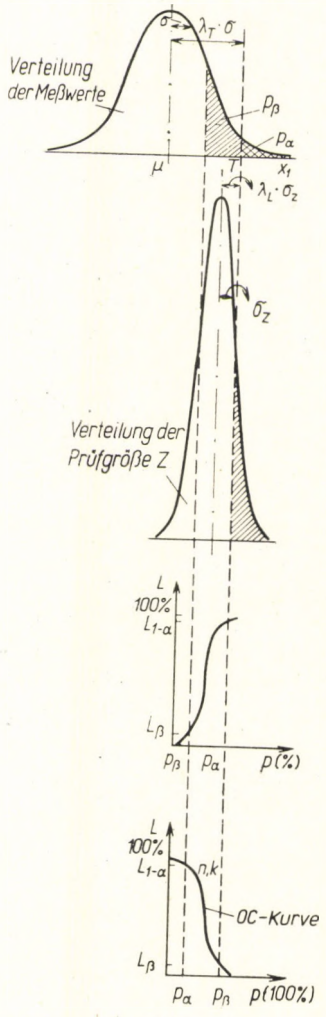


Abb. 1

Bei Stichprobenplänen der Form 2 wird die Größe λ_T mit Q bezeichnet.

Wie bereits erwähnt, treten im Falle der Stichprobenprüfung an die Stelle der Parameter μ und σ die entsprechenden Parameter der Stichproben, deren entsprechende Kombination das jeweilige Prüfsystem bestimmt.

Für die konstante Größe $\mu + \lambda_T \cdot \sigma$ existieren bei wiederholten Stichproben aus der gleichen Grundgesamtheit unzählige Größen, die sich um diese Ausgangsgröße gruppieren. Die Verteilung, die sich daraus ergibt, ist wiederum normal mit den Parametern μ_z und σ_z (μ_z wird bedingt durch die unterschiedliche Verteilungsdichte der Ausgangskurve links von $\mu + \lambda_T \cdot \sigma$ liegen — unterer Teil von Abb. 1).

Sowohl für μ_z als auch für σ_z besteht ein Zusammenhang zur Grundgesamtheit und es kommt darauf an, den funktionalen Zusammenhang zwischen den entsprechenden Parametern zu finden. Für σ_z, das ebenfalls in erster Linie vom Stichprobenumfang abhängig ist ($\sqrt{n}$-Gesetz) können Begrenzungen eingeführt werden in folgender Form:

1. Bei einem bestimmten λ_T-Wert p_β dürfen höchstens 10% der Lieferungen abgenommen werden;

2. Bei einem bestimmten λ_T-Wert p_α müssen mindestens 95% der Lieferungen angenommen werden.

Durch diese Festlegungen über p_α und p_β wird die OC-Kurve eindeutig bestimmt.

Die Parameter der einzelnen Stichprobensysteme lassen sich aus der Bestimmungsgleichung

$$\mu + \lambda_T \cdot \sigma = \mu_z + \lambda_L \cdot \sigma_z \tag{1}$$

ableiten.

Im Falle der $\bar{x}$-σ-Pläne gilt für die Prüfgröße:

$$z = \bar{x} + k \cdot \sigma$$

daraus folgt:

$$\mu_z = \mu + k \cdot \sigma$$

$$\sigma_z = \sigma/\sqrt{n}\,.$$

Werden diese Größen in die Gleichung (1) eingesetzt, erhalten wir

$$\lambda_L = \sqrt{n}(\lambda_T - k) . \tag{2}$$

Die Werte λ_L und λ_T können verschieden gedeutet werden. Einmal als Abszissenwert der normierten Gauß-Verteilung, zum anderen als die zu diesem Abszissenwert gehörigen Flächenanteile. (Dies ist der Zusammenhang, der in der Formel

$$F(u_0) = 1/\sqrt{2\pi} \cdot \int_{-\infty}^{u_0} e^{-u^2/2}\, du$$

zum Ausdruck kommt.)

λ_T entspricht der oberen Begrenzung des Integrals, λ_L der Fläche und damit im Fall der Stichprobenprüfung der Annahmewahrscheinlichkeit.

Um n und k aus (2) bestimmen zu können, müssen für die Annahmewahrscheinlichkeit L und die Toleranz T zwei Wertepaare gegeben sein. Sie werden zumeist den OC-Kurven der Attributpläne entnommen, um einmal die Gleichheit der OC-Kurven zu erhalten und zum anderen die Abhängigkeit vom Losumfang N zu umgehen.

Für $T = p_\alpha\,{}^0/_0 \rightarrow L = (1 - \alpha)$

für $T = p_\beta\,{}^0/_0 \rightarrow L = (\beta)$.

In (2) eingesetzt folgt für die $\bar{x}$-σ-Pläne nach den Formeln von Stange

$$n = \left(\frac{\lambda_{1-\alpha} + \lambda_{1-\beta}}{\lambda_{p\alpha} - \lambda_{p\beta}}\right)^2$$

$$k = \left(\frac{\lambda_{1-\alpha} + \lambda_{p\beta} + \lambda_{1-\beta} \cdot \lambda_{p\alpha}}{\lambda_{1-\alpha} + \lambda_{1-\beta}}\right) .$$

Wir haben für die $\tilde{x}$-σ- und für die $\tilde{x}$-QR_i-Pläne analoge Formeln entwickelt, wobei wir für die letzteren bedingt durch den begrenzten Definitionsbereich der einzelnen QR_i mittels der Methode der keinsten Quadrate eine Beziehung zwischen σ_{QRi} und dem σ der Grundgesamtheit ermittelt haben. In entsprechender Weise sind die Parameter der Variablenstichprobenpläne der Form 2 abzuleiten. Im Fall der $\bar{x}$-σ-Pläne erhalten wir den Qualitätsindex Q mit

$$Q = \left(\frac{T - \bar{x}}{\sigma}\right) \cdot v \tag{3}$$

der analog der Prüfgröße z der Form 1 wiederum für eine vorgegebene untere Toleranzgrenzen entsprechend zu berechnen ist. v ist ein Korrekturfaktor, der sich aus $v = \sqrt{n/n - 1}$ berechnet. Aus der Tafel der Normalverteilung können unter Einbeziehung der entsprechenden Korrekturfaktoren die M-Werte der Prüfpläne der Form 2 abgelesen werden. Sie wurden bei der Berechnung des MST 414 mit Hilfe der β- und der nichtzentralen t-Verteilung bestimmt.

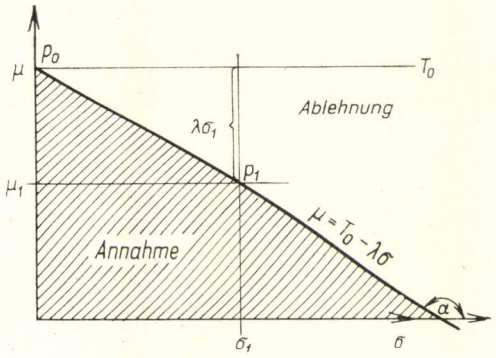


Abb. 2

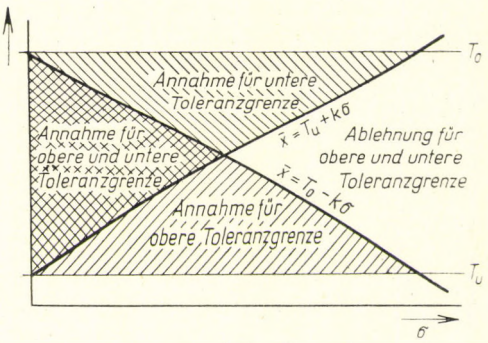


Abb. 3

Unabhängig davon, daß sich die Parameter der Variablenstichprobenpläne relativ einfach berechnen lassen und damit die Möglichkeit gegeben ist, in unkomplizierter Weise für spezielle Produktionsverhältnisse entsprechende Pläne zu entwickeln, bestehen für die praktische Anwendung Schwierigkeiten durch die mit der Anwendung dieser Pläne verbundenen Rechenoperation. So ist der Mittelwert zu bilden, das Streuungsmaß zu berechnen, es sind eine ganze Anzahl von Tabellen notwendig und obwohl die Operationen einfach sind, erfordern sie Zeit und vor allem qualifiziertes Personal.

Wir suchten deshalb nach einer Möglichkeit, diesen Aufwand zu beseiten oder zumindest zu reduzieren und fanden sie in der Konstruktion des sogenannten Annahmediagramms [1] (Abb. 2).

Das Annahmediagramm ist eine Umformung der rechnerischen Lösung auf eine graphische Entscheidung. Wird die Prüfgröße z gleich der Toleranz T gesetzt, so ergibt dies die allgemeine Form der Annahmegleichung mit

$$T_0 = \mu + \lambda_T \cdot \sigma$$

im Fall einer vorgegebenen oberen Toleranzgrenze, oder umgewandelt

$$\mu = T_0 - \lambda_T \cdot \sigma .$$

Dies ergibt im $\sigma - \mu$-Koordinatensystem eine Gerade. Durch Vorgabe der Toleranzgrenze und eines Wertepaares σ_1, μ_1 kann die zum entsprechenden Plan gehörige Annahmegerade eingezeichnet werden. Für Wertepaare, die bei der Stichprobenprüfung auf bzw. unterhalb dieser Geraden liegen, erfolgt Annahme, im anderen Fall Ablehnung.

Im Falle der Stichprobenprüfung treten anstelle der Parameter die Größen des gewählten Stichprobensystems, anstelle von λ_T der Faktor k. Dieser Faktor ist identisch mit dem Tangens des Winkels α, den die Gerade mit der σ-Achse bildet, so daß auch auf diese Weise die Konstruktion des Annahmediagrammes erfolgen kann.

Bei zweiseitig vorgegebenen Toleranzgrenzen, dem in der Praxis häufigeren Fall, ergeben sich für die Annahmeentscheidung folgende Bereiche (Abb. 3).

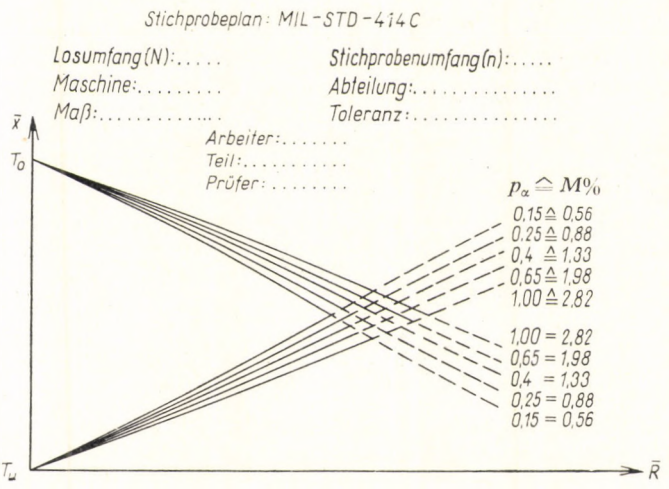


ABB. 4

Dieses Verfahren hat den Vorteil, daß bei gleicher Losgröße, gleichem P_α und gleichbleibenden Toleranzgrenzen — Bedingungen die für eine Serienfertigung zutreffen — das Diagramm nur einmal aufgebaut werden muß und dann immer wieder verwendet werden kann, solange sich die Bedingungen nicht ändern. Besonders wertvoll werden dabei die $\tilde{x}$-σ- und $\tilde{x}$-QR_l-Pläne, da hierbei jeglicher Rechenaufwand entfällt. Jedoch auch bei den anderen Systemen verringert sich der Aufwand gegenüber dem MST 414 erheblich, bei dem für zweiseitige Toleranzgrenzen die Annahmeentscheidung bei D-Plänen in max. 15, bei B-Plänen in max. 18 und bei C-Plänen in max. 15 Schritten fällt.

Eine ebenso wertvolle Erkenntnis ist, daß mit dem Annahmediagramm die Verbindung der Prüfpläne der Form 1 und Form 2 möglich wird. Aus der Identität des Qualitätsindex Q und des Toleranzfaktors λ_T im Fall der Grundgesamtheit, kann aus dem Annahmediagramm unmittelbar der zu erwartende Fehleranteil bestimmt werden.

Die Form des Annahmediagramms, die wir hierfür im Industriezweig verwenden zeigt Abb. 4.

Mit dieser Methode ist es möglich, gleichzeitig die Annahmeentscheidung, die Qualitätsgeschichte und die Bestimmung der Maschinenfunktionsgenauigkeit durchzuführen.

Mit dem Annahmediagramm wird es weiterhin möglich, die von mir einleitend als Punkt 3 und 4 genannten Forderungen zu realisieren, nämlich Prüfpläne mit reduziertem Aufwand zu berechnen. Es gibt bereits derartige Pläne. Einmal die Prüfpläne für reduzierte Prüfung des MST 414 zum anderen die Doppelstichprobenpläne für Variablenprüfung von Bowker und Goode. Die Pläne für reduzierte Prüfung des MST 414 genügen Forderung 4 nicht. Ihre OC-Kurve verläuft wesentlich flacher und gibt damit für schlechte Qualitäts eine zu hohe Annahmewahrscheinlichkeit. Die

Pläne von Bowker and Goode lehnten wir für unsere Zwecke ab, da einfach- und Doppelpläne unterschiedliche OC-Kurven besitzen.

In dem Bestreben, für qualitativ gute Fertigungsabschnitte eine rationellere Prüfung zu entwickeln, setzten wir bei den Doppelstichproben $n_1 = n/2$ und bei den Mehrfachstichprobenplänen $n_1 = n/3$ fest. Der theoretische Aufbau der Doppelstichprobenpläne für Variablenprüfung erfolgt von den gleichen Voraussetzungen, wie die Berechnung der Einfachstichprobenpläne, d. h. von der Annahme einer hinreichend normalen Verteilung der Parameter der einzelnen Stichprobensysteme. Da diese Annahme speziell im Bereich der kleinen Stichprobenumfänge zu Fehlern führen kann, wurden Doppel- und Mehrfachpläne nur für jene Losgrößen berechnet, bei denen der Umfang der ersten Stichprobe im Bereich der Stichprobenumfänge der Einfachstichprobenpläne liegt. Hiermit wird jedoch faktisch keine Einschränkung vorgenommen.

Als weitere Voraussetzung stellten wir die Bedingung, daß die Gesamt-Annahmewahrscheinlichkeit von Doppel- und Mehrfachplänen mit der der Einfachstichprobenpläne identisch ist.

Für die Annahmeentscheidung für Doppelpläne können damit folgende Fälle auftreten:

Annahme nach der ersten Stichprobe erfolgt, wenn

$$z_1 \leqq T$$

Annahme nach der zweiten Stichprobe erfolgt, wenn

$$z_1 > T$$

aber

$$z \leqq T$$

Ablehnung erfolgt, wenn

$$z_1 > T$$

und

$$z > T$$

Für Mehrfachstichprobenpläne gilt dieses Schema in analog erweiterter Weise.

Für die praktische Anwendung setzen wir wiederum das Annahmediagramm voraus, da eine rechnerische Entscheidung durch die Vielzahl der notwendigen Entscheidungsschritte nicht vertretbar ist. Das Diagramm zeigt Abb. 5.

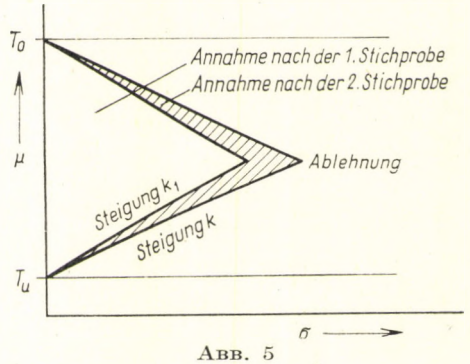


Abb. 5

Der theoretische Aufbau der Bestimmungsgleichung der Kenngrößen für Doppelstichprobenpläne entspricht im wesentlichen den der Einfachpläne. Der Stichprobenumfang wurde von vornherein festgelegt, so daß lediglich noch die k_1-Werte zu bestimmen waren. Zur Bestimmung dieser Größen mußten Festlegungen über das Verhältnis der Annahmewahrscheinlichkeit der einzelnen Stichproben zueinander getroffen werden. Die dazu notwendigen Tabellen berechneten wir mit Hilfe des in der DDR produzierten elektronischen Kleinrechners 'Cellatron' SER 2 b.

LITERATUR

1. *Industrial Quality Control.* April 1962.

ON THE OC-FUNCTION OF THE ACCEPTANCE PLAN DESIGNED FOR A PÓLYA PROCESS OF DEFECTIVE ITEMS

by

G. TUSNÁDY

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary

1. INTRODUCTION

In the application of the methods of mathematical statistics to the design of acceptance plans the fact that the production process is a more or less complicated stochastic process has not been taken into account until recently. The control chart methods of the production control and the acceptance plans as well are based in general on the assumption that the quality characteristics of subsequently produced products are independent of each other. In order to obtain a better agreement of theory and practice the theory of stochastic processes has to be applied.

This conclusion has been drawn in a paper of I. Vincze [1] on the OC-functions of acceptance plans designed for a Pólya process of the defective items.

Let us now consider a discrete production process, in which the quality of the products is either good or defective. We assign to each product either the number 0 or 1, according to whether the product is defective or good; let us denote by ε_k the number corresponding to the k-th product. Then the so-called (n, c)-plan is an acceptance device, according to which we accept the lot if in a sample of n elements there are less than c defective pieces* and reject it otherwise. If p is the probability of the variable ε_k being equal to 1, then the probability of acceptance depends besides of p on the stochastic connection of the variables ε_k. It seems that in case the variables ε_k are positively correlated, less information is obtained for the probability p, but more information for the number of defective pieces in the lot than in the case of independence. Our following theorem is a first step in the direction of giving a precise meaning to this heuristic statement for the case when $\{\varepsilon_k\}$ follows a Pólya process.

2. THE OC-FUNCTION OF THE PÓLYA-PROCESS

Let ε_k be 0 or 1 ($k = 1, 2, \ldots$) and $\xi_k = \sum_{i=1}^{k} \varepsilon_i$. Then $\{\varepsilon_k\}$ is a Pólya-process, if

$$\mathbf{P}(\varepsilon_{k+1} = 1 \mid \varepsilon_1, \ldots, \varepsilon_k) = \frac{p + q\xi_k}{1 + qk} \tag{1}$$

* This is a slight modification of the definition used in general, according to which the lot is accepted even if in the sample there are c defectives.

i.e. if the conditional probability of a $(k+1)$-st item being defective depends only on ξ_k the number of defectives in the first k items and is a monotonously increasing function of this number. We assume that the second parameter of the process is positive: $q > 0$; in case $q = 0$ the independent process with binomial distribution is obtained.

The OC-function gives the probability of acceptance as a function of the parameter p. As the ε_k-s are equivalent (exchangeable) random variables, we may assume that the sample consists of the first n items of the process, and thus

$$L(p \mid q, n, c) = \mathbf{P}(\xi_n < c \mid p, q) \tag{2}$$

The selectivity of the acceptance plan depends on the slope of this function; it is natural to expect that to a larger parameter q there corresponds a flatter OC-function. This is the conclusion of the following theorem.

THEOREM. *Let* $0 \leq q < Q$, $n_1 \geq n_2$, $1 \leq c_1 \leq n_1$, $1 \leq c_2 \leq n_2$; *then there is a* $p_0 \in [0, 1]$ *such that*

$$\begin{aligned} L(p \mid q, n_1, c_1) &> L(p \mid Q, n_2, c_2) \quad \text{if} \quad 0 < p < p_0 \\ L(p \mid q, n_1, c_1) &< L(p \mid Q, n_2, c_2) \quad \text{if} \quad p_0 < p < 1 \end{aligned} \tag{3}$$

i.e. the graph of the function $L(p|Q, n_2, c_2)$ *intersects the graph of the function* $L(p|q, n_1, c_1)$ *only once, from below.*

For the case $c_1 = 1$ or n_1 the following inequalities hold:

$$\begin{aligned} L(p \mid q, n_1, 1) &< L(p \mid Q, n_2, c_2) \quad \text{if} \quad 0 < p < 1 \\ L(p \mid q, n_1, n_1) &> L(p \mid Q, n_2, c_2) \quad \text{if} \quad 0 < p < 1. \end{aligned}$$

COROLLARY. If we find to the numbers $1 \leq c_1 \leq n_1$, $0 < p_1 < p_2 < 1$, $Q > 0$ numbers n_2, c_2 such that the following inequalities are fulfilled:

$$\begin{aligned} L(p_1 \mid 0, n_1, c_1) &\leq L(p_1 \mid Q, n_2, c_2) \\ L(p_2 \mid 0, n_1, c_1) &\geq L(p_2 \mid Q, n_2, c_2) \end{aligned}$$

then $n_2 > n_1$.

Indeed, assume indirectly that $n_2 \leq n_1$. Then in virtue of our theorem there is a p_0 such that (3) holds with $q = 0$; this would imply $p_2 < p_0$, $p_1 > p_0$, which is a contradiction.

In this respect our theorem is a generalisation of a conclusion of paper [1] concerning the plans $(n, 0)$ and for the general case (n, c). We remark here, that for small values of the parameter q a proof of our theorem has been given by E. Vas (unpublished result).

PROOF. Let us consider the OC functions for a fixed q. For fixed values of n and p the function $L(p|q,n,c)$ increases as the parameter c increases; for fixed n and c it decreases if the variable p increases, and for $p = 0$ or 1 the value of the function L is 1 or 0, resp. Thus for fixed n the functions

$$L_c(p) = L(p|q,n,c) \qquad (c = 1, \ldots, n) \tag{4}$$

subdivide the unit square of the plane (p, L) into $(n+1)$ parts. Let us put $L_0(p) \equiv 0$ and $L_{n+1}(p) \equiv 1$, and let us form the following functions:

$$L(p) = \lambda L_c(p) + (1-\lambda) L_{c+1}(p) \tag{5}$$

$$c = 0, 1, \ldots, n\,;\ 0 < \lambda \leq 1$$

then the unit square of the plane (p, L) is covered by the functions $L(p)$ in such a way that to any

$$0 < p_0 < 1\,;\ 0 < L_0 < 1 \tag{6}$$

there is one and only one function $L(p)$ with $L(p_0) = L_0$. Let us denote by $\mathscr{L}(n)$ the class of the functions $L(p)$. By the aid of the class $\mathscr{L}(n)$ we define a new function $D(L|n, p_0)$, assigning to each point $L_0 \in [0, 1]$ the slope at p_0 of that function $L(p)$ for which $L(p_0) = L_0$ i.e.

$$D(L_0 \,|\, n, p_0) = L'(p_0) \quad \text{where} \quad L(p_0) = L_0 \tag{7}$$

The function $D(L|n, p)$ is trivially continuous, and piecewise linear. As a first step we prove that $D(L|n, p)$ is convex from below and for fixed (p, L) a decreasing function of the parameter n. We need the following recursive formulae:

$$L(p \,|\, q, n+1, k+1) = \frac{p+kq}{1+nq}\, L(p \,|\, q, n, k) + \\ + \left(1 - \frac{p+kq}{1+nq}\right) L(p \,|\, q, n, k+1). \tag{8}$$

Indeed,

$$L(p \,|\, q, n+1, k+1) = \mathbf{P}(\xi_{n+1} < k+1) = \sum_{i=0}^{n} \mathbf{P}(\xi_n = i)\, \mathbf{P}(\xi_{n+1} < k+1 \,|\, \xi_n = i) =$$

$$= \sum_{i=0}^{k-1} \mathbf{P}(\xi_n = i) + \mathbf{P}(\xi_n = k)\, \mathbf{P}(\xi_{n+1} < k+1 \,|\, \xi_n = k) =$$

$$= L(p \,|\, q, n, k) + \mathbf{P}(\xi_n = k) \left(1 - \frac{p+kq}{1+nq}\right).$$

Furthermore

$$L(p \,|\, q, n, k) = \frac{n+1-k}{n+1} L(p \,|\, q, n+1, k) + \frac{k}{n+1} L(p \,|\, q, n+1, k+1). \tag{9}$$

In fact

$$L(p \,|\, q, n, k) = \mathbf{P}(\xi_n < k) = \sum_{i=0}^{n+1} \mathbf{P}(\xi_{n+1} = i)\, \mathbf{P}(\xi_n < k \,|\, \xi_{n+1} = i) =$$

$$= \sum_{i=0}^{k-1} \mathbf{P}(\xi_{n+1} = i) + \mathbf{P}(\xi_n < k \,|\, \xi_{n+1} = k)\, \mathbf{P}(\xi_{n+1} = k) =$$

$$= L(p \,|\, q, n+1, k) + \frac{k}{n+1} \mathbf{P}(\xi_{n+1} = k).$$

We remark that (8) is a special case of the more general formula

$$L(p\,|\,q,n+1,k+1) = \mathbf{P}(\varepsilon_{n+1}=1\,|\,\xi_n=k)\,L(p\,|\,q,n,k) + \\ + \mathbf{P}(\varepsilon_{n+1}=0\,|\,\xi_n=k)\,L(p\,|\,q,n,k+1) \tag{8*}$$

valid for any process with two possible states 0 and 1, and formula (9) is valid for any sequence of equivalent variables with two possible values 0 and 1.

On the basis of the formulae (8) and (9) we prove our assertion by induction for the parameter n. Since the graph of the function $D(L|1,\ p)$ consists of two segments, and

$$D(0\,|\,1,p) = D(1\,|\,1,p) = 0\,; \qquad D(L\,|\,1,p) \leq 0\,,$$

$D(L|1,p)$ is a convex function. Let us assume that the function $D(L|n,\ p)$ is convex. The formula (8) implies that the peaks of the function $D(L|n+1,p)$ are below the graph of the function $D(L|n,\ p)$. Indeed, let $f(p)$ and $g(p)$ be two functions for which

$$f(0) = g(0) = 1\,; \qquad f(1) = g(1) = 0\,; \\ f(p) < g(p) \quad \text{for} \quad p\in(0,1) \tag{10}$$

and let the function $h(p)$ be defined as

$$h(p) = \lambda(p)f(p) + [1-\lambda(p)]g(p) \tag{11}$$

where $\lambda(p)$ is a monotonously increasing function, and $0 < \lambda(p) < 1$; then for $p \in (0,\ 1)$ we obtain

$$h'(p) = \lambda'(p)[f(p)-g(p)] + \lambda(p)f'(p) + [1-\lambda(p)]g'(p)$$

where the first term of the right side is negative, thus at a point p_0 the slope of the function $h(p)$ is less than the slope of that linear combination of the functions $f(p)$, $g(p)$ which intersects $h(p)$ at p_0. In our case for the functions

$$f(p) = L(p\,|\,q,n,k) \\ g(p) = L(p\,|\,q,n,k+1) \\ \lambda(p) = \frac{p+kq}{1+nq} \tag{12}$$

the conditions (10), (11) are valid, hence the peaks of the function $D(L|n+1,p)$ are below the graph of the function $D(L|n,\ p)$.

From formula (9) there follows that the functions $L(p|q,\ n,\ c)$ $(c = 0,\ 1,\ \ldots,\ n+1)$ which span the class $\mathscr{L}(n)$ are linear combinations of the functions $L(p|q,\ n+1,\ c)$, i.e. they belong to the class $\mathscr{L}(n+1)$. Thus the peaks of the function $D(L|n,\ p)$ are on the linear segments of the graph of the function $D(L|n+1,\ p)$, and as we have seen, the peaks of $D(L|n+1,\ p)$ are below the graph of the function $D(L|n,\ p)$. This means that $D(L|n+1,\ p) \leq$ $\leq D(L|n,\ p)$ and the convexity of $D(L|n,\ p)$ implies that of $D(L|n+1,\ p)$.

We shall say that a process $\{\varepsilon_k\}$ satisfies the convexity condition if its functions $D(L|n, p)$ are convex in the variable L and are monotonous in the parameter n.

In the next step of our proof we shall need the following concept: the process $\{\eta_k\}$ is a secondary process of the Pólya type of the process $\{\varepsilon_k\}$ if $\eta_k = \varepsilon_{\nu(k)}$ where the process $\nu(k)$ is an integer valued process in which from the state i one-step transitions to the states $0 \leq j \leq i+1$ are only possible. Though the term 'Pólya type' is justified only in the case when the transition possibilities of the process $\nu(k)$ are according to formula (13), we shall use this name for the general case, too.

We assign to the process $\{\varepsilon_k\}$ and $\{\eta_k\}$ the function classes $\mathscr{L}_\varepsilon(m)$ and $\mathscr{L}_\eta(n)$ and the functions $D_\varepsilon(L|m, p)$, $D_\eta(L|n, p)$ resp. From our definition it follows that the OC function $L(p|\eta, n, k)$ of the secondary process is a linear combination of the OC-functions $L(p|\varepsilon, m, k)$ of the process $\{\varepsilon_k\}$ with parameters $m \leq n$. Thus the vector of the peaks of the function $D_\eta(L|n, p)$ is a linear combination of vectors of the peaks of the functions $D_\varepsilon(L|m, p)$ with $m \leq n$. If the process $\{\varepsilon_k\}$ satisfies the convexity condition, we have $D_\eta(L|n, p) \leq D_\varepsilon(L|n, p)$ and each member of the class $\mathscr{L}_\eta(n)$ intersects each member of the class $\mathscr{L}_\varepsilon(n)$ from below.

Utilizing this result, the proof of our theorem will be complete if we prove that a Pólya process with parameter Q can be represented as a secondary process $\{\eta_k\}$ of Pólya type of the process $\{\varepsilon_k\}$ with parameter q. Let us write

$$\nu(1) = 1$$

$$\mathbf{P}(\nu(n+1) = i \mid \nu(1), \ldots, \nu(n)) = \begin{cases} \dfrac{\beta(i)\,Q - q}{1 + nQ} & \text{if } 1 \leq i \leq M_n \\ \dfrac{1 + M_n q}{1 + nQ} & \text{if } i = M_n + 1 \end{cases}$$

where $M_n = \max\{\nu(1), \ldots, \nu(n)\}$

$\beta(i) = \{\text{the number of } j\text{'-s, for which } \nu(j) = i,\ 1 \leq j \leq n\}$

We prove that the process $\eta_k = \varepsilon_{\nu(k)}$ is a Pólya process with parameter Q; to this end we prove by induction the following relation:

$$\mathbf{P}(\eta_{j+1} \mid \eta_1, \ldots, \eta_j) = \frac{p + Q \sum_{i=1}^{j} \eta_i}{1 + jq} \tag{14}$$

For $j = 0$ this is trivially true. Let us suppose that it holds for all $j < n$, where $n \geq 1$. We have:

$$\mathbf{P}(\eta_{n+1} = 1 \mid \eta_1, \ldots, \eta_n) =$$

$$\sum_{k=1}^{n} \mathbf{P}(\nu(n) = k \mid \eta_1, \ldots, \eta_n)\mathbf{P}(\eta_{n+1} = 1 \mid \eta_1, \ldots, \eta_n;\ \nu(n) = k). \tag{15}$$

Let us put

$$p_m = \mathbf{P}(\xi_k = m \mid \eta_1, \ldots, \eta_n; \nu(n) = k) \quad \text{where} \quad \xi_k = \sum_{i=1}^{k} \varepsilon_i . \tag{16}$$

By our assumption (14) the joint distribution of the variables $(\eta_1, \eta_2, \ldots, \eta_n)$ is the same as that of the first n variables of a Pólya process with parameter Q. Thus the above conditional distribution of the variable ξ_k is a hypergeometric distribution:

$$p_m = \frac{\binom{\xi_n}{m}\binom{n-\xi_n}{k-m}}{\binom{n}{k}} \qquad \text{where} \qquad \xi_n = \sum_{i=1}^{n} \eta_i . \tag{17}$$

Substituting the conditional probability p_m in the second term of formula (15), we obtain

$$\mathbf{P}(\eta_{n+1} = 1 \mid \eta_1, \ldots, \eta_n; \nu(n) = k) =$$

$$= \sum_{m=1}^{k} p_m \mathbf{P}(\eta_{n+1} = 1 \mid \eta_1, \ldots, \eta_n; \nu(n) = k; \xi_k = m) =$$

$$= \sum_{m=1}^{k} p_m \left(\frac{1+kq}{1+nQ} \cdot \frac{p+mq}{1+kq} + \frac{nQ-kq}{1+nQ} \frac{m}{k} \right) =$$

$$= \sum_{m=1}^{k} p_m \frac{p + m\frac{n}{k}Q}{1+nQ} = \frac{p + \frac{n}{k}Q \sum_{m=1}^{k} m p_m}{1+nQ}$$

where $\sum_{m=1}^{k} m p_m$ is the conditional expectation of ξ_k, thus by (17)

$$\sum_{m=1}^{k} m p_m = \frac{k}{n} \xi_n \tag{18}$$

and by (15) it follows that (14) holds for $j = n$.

The proof of our theorem is completed.

REMARK. The fact that the distribution of the variables η_k of Pólya type with parameter Q is for $q = 0$ a mixture of binomial distributions can be regarded as a special case of de Finetti's theorem [2].

3. REMARKS AND PROBLEMS

We return now to some problems mentioned in the introduction. Let us suppose that in two factories the percent-defectives p are equal, but the Pólya-parameters q and Q differ: $q < Q$. From our theorem there does not follow in general the necessity of a larger sample for the factory with para-

meter Q, but it does in case we wish to 'fit' the OC-function in two given points—in the sense as this was described in the Corollary. But we believe that this 'fitting' of the OC-functions is not a natural requirement. This can be seen by the fact that the expected number of defectives in a lot of given size is a monotonously decreasing function of the parameter Q. Let us put $Q = \infty$, i.e. the quality of each piece will be determined by the first piece: if the first piece is good, all subsequent pieces will be good and if it is defective, all products will be defective. These events may occur with probabilities $(1 - p)$ and p resp. In this case the OC-function coincides with the line $L = 1 - p$, i.e. with the worst OC-function for the binomial case, corresponding to the $n = 1$. Nevertheless in an accepted lot all the products are good and in a rejected lot they are all defective.

This example shows that in case of dependent sample elements the OC-functions are not always a suitable basis for deciding which of two plans gives a better protection against faulty decisions. Thus the question arises, what further types of characteristics could be used to approach this problem from other directions.

REFERENCES

1. Vincze, I. A tömeggyártás minőségellenőrzésének matematikai statisztikai módszereiről. (In Hungarian) *MTA III. Oszt. Közl.* IV. **3,** (1964).
2. De Finetti, B. La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré.* **7,** 1–68 (1937).

ON THE POWER OF THE KOLMOGOROV–SMIRNOV TWO-SAMPLE TEST AND RELATED NON-PARAMETRIC TESTS*

by

I. VINCZE

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary

INTRODUCTION

In the present paper we shall consider the Kolmogorov–Smirnov two sample test in the case of equal sample sizes. In the first part the power function of this test is given for the following kind of alternatives:

Let ξ and η be random variables with continuous distribution functions $F(x)$ and $G(x)$ resp. Let the null-hypothesis be:

$$H_0 : G(x) \equiv F(x)$$

and the alternative

$$H_1 : G(x) \equiv \psi\big(F(x)\big),$$

where $\psi(y)$ is a piecewise linear distribution function in the interval (0,1), $\psi(y) \not\equiv y$. The distribution functions of this kind are everywhere dense in the set of continuous alternatives. As an application of our expression for the power function we shall determine numerically the second kind error of the one sided Kolmogorov–Smirnov test in the case of a simple alternative for which $\psi(y)$ contains two linear sections. The results show that in using this test we have to take care in the case of small samples: The second kind error may be relatively high in the case of sample sizes 20—50.

In the second part we shall deal with the following question: For diminishing the difference in efficiency between parametric and nonparametric tests, the author has in his papers [7, 9] proposed the use of a pair of statistics, instead of one statistic. In consequence of the Neyman–Pearson lemma, this results for a given simple alternative in a better test than the one based on either single test-statistic. We apply the Kolmogorov–Smirnov statistic, i.e. the maximum deviation of the two empirical distribution functions as the first statistic, which ensures the asymptotic efficiency of the test, and we add to this the first maximum-index (defined below) as the second statistic. Constructing the power function of the test based on this pair of statistics for the same alternative as mentioned before we may observe the increase in the efficiency compared with the Kolmogorov—Smirnov test. — We shall add several remarks to the problems treated.

* The material of the author's lecture was published in [11] in Hungarian; the present text contains some results of [10], too.

1. THE POWER FUNCTION IN THE GNEDENKO–KOROLJUK CASE

1. Referring to the alternative mentioned in the introduction we may restrict ourselves to the following hypotheses:

$$H_0: G(x) \equiv F(x) = \begin{cases} 0, & \text{if } x \leqq 0 \\ x, & \text{if } 0 \leqq x \leqq 1 \\ 1, & \text{if } 1 \leqq x \end{cases}$$

$$H_1^{(r)}: \begin{cases} \text{F(x) uniformly distributed in (0,1)} \\ G'(x) = g(x) = \begin{cases} 0, & \text{if } x \leqq 0 \\ g_i, & \text{if } z_{i-1} \leqq x < z_i, \; i = 1, 2, \ldots r \\ 0, & \text{if } 1 \leqq x, \end{cases} \end{cases}$$

where $0 = z_0 < z_1 < z_2 < \ldots < z_r = 1$. For the non-negative values g_i the following relation holds: $\int_0^1 g(x)dx = \sum_{i=1}^{r} g_i(z_i - z_{i-1}) = 1$.

Let further $\xi_1, \xi_2, \ldots \xi_n$ and $\eta_1, \eta_2, \ldots \eta_n$ be independent observations on ξ and η resp. with the corresponding empirical distribution functions $F_n(x)$ and $G_n(x)$ resp. In deciding between H_0 and $H_1^{(r)}$ we shall use the statistic

$$D_{n,n} = \max_{(x)} | F_n(x) - G_n(x) |$$

or in the case of one-sided alternatives

$$D_{n,n}^{+} = \max_{(x)} \big(F_n(x) - G_n(x)\big).$$

The α-size critical region of the test will contain those samples for which the events $\{D_{n,n} \geqq d_\alpha\}$ or $\{D_{n,n}^{+} \geqq d_\alpha^{+}\}$ hold, where the positive numbers d_α and d_α^{+} are defined in the following way:

$$P(D_{n,n} < d_\alpha \,|\, H_0) = 1 - \alpha,$$

$$P(D_{n,n}^{+} < d_\alpha^{+} \,|\, H_0) = 1 - \alpha.$$

The power functions will be

$$K_n(H_1^{(r)}, \alpha) = P(D_{n,n} \geqq \alpha \,|\, H_1^{(r)})$$

and

$$K_n^{+}(H_1^{(r)}, \alpha) = P(D_{n,n}^{+} \geqq \alpha \,|\, H_1^{(r)}).$$

Notation for the probability of a further event will now be introduced. We denote the elements of the ordered samples by

$$\xi_1^* < \xi_2^* < \ldots < \xi_n^*$$

and

$$\eta_1^* < \eta_2^* < \ldots < \eta_n^*$$

resp. Let further

$$\tau_1^* < \tau_2^* < \ldots < \tau_{2n}^* \tag{1}$$

be the union of the two samples in order of magnitude. We define now the random variables for $i = 1, 2, \ldots 2n$

$$\vartheta_i = \begin{cases} +1, & \text{if } \tau_i^* = \xi_j, \\ -1, & \text{if } \tau_i^* = \eta_2 \end{cases}$$

for some j and l. With the usual notation

$$s_0 = 0, s_i = \vartheta_1 + \vartheta_2 + \ldots + \vartheta_i (i = 1, 2, \ldots 2n), s_{2n} = 0.$$

The points (i, s_i) in a plane rectangular coordinate system give the path of a random walk starting at the origin and returning after $2n$ steps to the point $(2n, 0)$. As it is known and can be seen to be very simple, the relations

$$D_{n,n} = \frac{1}{n} \max_{(i)} |s_i| \quad \text{and} \quad D_{n,n}^+ = \frac{1}{n} \max_{(i)} s_i$$

hold.

It is a simple fact too, that under the validity of H_0 in the sequence $(\vartheta_1, \vartheta_2, \ldots \vartheta_{2n})$ each array of $n(+1)$'s and $n(-1)$'s have the same probability, i.e. $\binom{2n}{n}^{-1}$.

Let us denote by the vectors $\{\nu\} = \{\nu_1, \nu_2, \ldots \nu_r\}$ and $\{\mu\} = \{\mu_1, \mu_2, \ldots \mu_r\}$ the events that are contained in the interval (z_{j-1}, z_j) from the n ξ_i-s just number ν_j and independently out of the n η_e-s just μ_j $(j = 1, 2, \ldots r)$. The probabilities of these events are clearly

$$\mathbf{P}(\{\nu\} \mid H_0) = \mathbf{P}(\{\nu\} \mid H_1^{(r)}) = n! \prod_{j=1}^{r} \frac{(z_j - z_{j-1})^{\nu_j}}{\nu_j!},$$

$$\mathbf{P}(\{\mu\} \mid H_0) = n! \prod_{j=1}^{r} \frac{(z_j - z_{j-1})^{\mu_j}}{\mu_j!},$$

$$\mathbf{P}(\{\mu\} \mid H_1^{(r)}) = n! \prod_{j=1}^{r} \frac{[(z_j - z_{j-1})\, g_j]^{\mu_j}}{\mu_j!}.$$

Further for $i = 0,1$,

$$\mathbf{P}(\{\nu\}\{\mu\} \mid H_i) = \mathbf{P}(\{\nu\} \mid H_i) \,.\, \mathbf{P}(\{\mu\} \mid H_i) \,.$$

Denoting now by n_j and m_j for $j = 1, 2, \ldots r$ the partial sums $\nu_1 + \nu_2 + \ldots + \nu_j$ and $\mu_1 + \mu_2 + \ldots + \mu_j$ resp. $(n_0 = m_0 = 0)$, it holds the following:

THEOREM: *The α-size two sample Kolmogorov–Smirnov test has the power function under $H_1^{(r)}$ against H_0*

$$K_n(H_1^{(r)}, \alpha) = 1 - (n!)^2 \sum_{(\nu)}^{*} \sum_{(\mu)}^{*} p_{jk^+} \frac{(z_j - z_{j-1})^{\nu_j + \mu_j} g_j^{\mu_j}}{\nu_j! \mu_j!}$$

with

$$p_{jk^+} = 1 - \frac{\binom{\nu_j + \mu_j}{\mu_j + k^+ - n_{j-1} + m_{j-1}}}{\binom{\nu_j + \mu_j}{\nu_j}},$$

when $H_1^{(r)}$ is a one sided alternative $G(x) < F(x)$, further $k^+ = k^+(\alpha, n) = nd_\alpha^+$ and

$$p_{jk} = \frac{1}{\binom{\nu_j + \mu_j}{\nu_j}} \sum_{s=-\infty}^{\infty} \left[\binom{\nu_j + \mu_j}{\nu_j} - \binom{\nu_j + \mu_j}{\nu_j + n_{j-1} - m_{j-1} + (2s+1)k} \right]$$

in the case of a two sided alternative with $k = k(\alpha, n) = nd_\alpha$.

Here $\sum_{(\nu)}^{*}$ and $\sum_{(\mu)}^{*}$ denote r-summation for all possible ν and μ vectors with $0 \leqq \nu_j, \mu_j \leqq n$ and $\sum_{j=1}^{r} \nu_j = \sum_{j=1}^{r} \mu_j = n$ and with the further restriction $n_j - m_j < k^+$ in the one sided case and $| n_j - m_j | < k$ in the two sided case, $j = 1, 2, \ldots r$.

This theorem is the consequence of the following simple

LEMMA. Let ξ and η be continuous random variables in the interval (0,1), (the density function of which are (not necessarily equal) constants in a subinterval (a, b). Let $\xi_1, \xi_2, \ldots \xi_\nu$ and $\eta_1, \eta_2, \ldots \eta_\mu$ be independent observations on ξ, and η resp. falling in the interval (a, b). Under this condition all possible arrays of the mentioned $\nu + \mu$ sample elements have the common probability $\binom{\nu+\mu}{\nu}^{-1}$.

In consequence of this lemma under the condition $\{\nu, \mu\}$ in each interval (z_{j-1}, z_j) the probabilities of the events $\{s_i < k^+\}$ and $\{|s_i| < k\}$ can be determined when H_0 is true.

Let us turn now to the determination of the probability $1 - K_n(H_1^{(r)}, \alpha) = \mathbf{P}(nD_{n,n}^+ < k \mid H_1^{(r)})$. We shall calculate the probability under each condition $\{\nu, \mu\}$ and multiply this by $\mathbf{P}(\{\nu, \mu\} \mid H_1^{(r)})$; then we have to sum, over all ν and μ (the same considerations are valid for the two-sided case).

The event $\{nD_{n,n}^+ < k^+\} = \{s_i < k^+\}$ means that the mentioned random walk does not reach the height k^+. But under our condition the random walk falls into parts with division points $(i_j = n_j + m_j, s_{i_j} = n_j -$

$- m_j)$ $(j = 1, 2, \ldots r - 1)$ and according to the Markov-property of the random walk the relation

$$\mathbf{P}\left(D_{n,n}^{+} < \frac{k^+}{n} \middle| \nu, \mu, H_1^{(r)}\right) =$$

$$= \prod_{j=1}^{r} P(s_i < k^+ \quad \text{for} \quad n_{j-1} + m_{j-1} < i < n_j + m_j \,|\, \mu, \nu, H_1^{(r)})$$

holds. Taking into account the elementary formula

$$\mathbf{P}(s_i < k^+ \quad \text{for} \quad n_{j-1} + m_{j-1} < i < n_j + m_j \,|\, \nu, \mu, H_1^{(r)}) =$$

$$= 1 - \frac{\binom{\nu_j + \mu_j}{\mu_j + k^+ - n_{j-1} + m_{j-1}}}{\binom{\nu_j + \mu_j}{\nu_j}}$$

and the relations given above for $P(\{\nu, \mu\} \mid H_1^{(r)})$ we come to the proof of our theorem.

2. We shall treat now the special case $r = 2$, i.e. the case when the alternative consists of two linear sections. When the notations $z_1 = z(z_0 = 0, z_2 = 1)$ and $g_1 = g(0 \leq g < 1)$ are used, then $g_2 = \dfrac{1 - gz}{1 - z}$, further $\nu_1 = \nu$, $\nu_2 = n - \nu$, $\mu_1 = \mu$, $\mu_2 = n - \mu$. As under this condition the alternative $H_1^{(2)}$ is one-sided, we may apply the one sided test. The combinatorial part of the probabilities reduce now to the formulas

$$\mathbf{P}(s_i < k^+ \quad \text{for} \quad 0 < i < \nu + \mu \,|\, H_1^{(2)}) = 1 - \frac{\binom{\nu + \mu}{\nu - k^+}}{\binom{\nu + \mu}{\nu}}$$

$$\mathbf{P}(s_i < k^+, \quad \text{for} \quad \nu + \mu < i < 2n \,|\, H_1^{(2)}) = 1 - \frac{\binom{2n - \nu - \mu}{n - \nu + k^+}}{\binom{2n - \nu - \mu}{n - \nu}}.$$

As mentioned $\nu + \mu < k$ must hold. Denoting the product of these two probabilities by $C_{\nu,\mu}^{n,k}$ we come to the following form of the power function:

$$K_n(H_1^{(2)}, \alpha) = 1 - \sum_{(\nu)}^{*} \sum_{(\mu)}^{*} \binom{n}{\nu} z^{\nu}(1 - z)^{n-\nu} \binom{n}{\mu} (gz)^{\mu} (1 - gz)^{n-\mu} C_{\nu,\mu}^{n,k},$$

where the summation holds for $0 \leqq \nu \leqq n$, $0 \leqq \mu \leqq n$ under the condition $\nu - \mu < k^+$.

From this formula the interesting consequence follows: if $n \to \infty$, then $k^+ \sim y\sqrt{2n}$ at one hand and $\nu \sim nz + \lambda\sqrt{nz(1-z)}$, $\mu \sim nzg + \lambda'\sqrt{nzg(1-zg)}$ on the other hand, which means that the requirement

$$\nu - \mu \sim nz(1-g) + \lambda\sqrt{nz(1-z)} - \lambda'\sqrt{nzg(1-zg)} < y\sqrt{2n}$$

can be satisfied only for λ-s, which tend to infinity. Therefore in determining the power, the normal approximation of the binomial terms is not suitable. A similar remark was made by J. Rosenblatt [6].

Turning now to the case $g = 0$, in which $\max(F(x) - G(x)) = F(z) - G(z) = z$ and $\mathbf{P}(\mu = 0 \mid H_1^{(2)}) = 1$ we obtain the following simple form of the power function:

$$K_n(H_1^{(2)}, \alpha) = 1 - \sum_{\nu=0}^{k-1} \binom{n}{\nu} z^\nu (1-z)^{n-\nu} \left(1 - \frac{\binom{2n-\nu}{n-\nu+k^+}}{\binom{2n-\nu}{n-\nu}} \right).$$

A slight modification leads to the expession:

$$K_n(H_1^{(2)}, \alpha) = 1 - B(1-z; n-k^+ + 1, k^+) + \frac{\alpha}{(1-z)^k} B(1-z; n+1, k^+),$$

where B is the Beta-function

$$B(z; p, q) = \frac{\int_0^z t^{p-1}(1-t)^{q-1}\,dt}{\int_0^1 t^{p-1}(1-t)^{q-1}\,dt}.$$

This formula is suitable for immediate computation. We have chosen this simple alternative for we need the power function in the more complicated two statistic case.

Now we give some numerical values for small samples using (n, k^+) pairs for which α is near 10% and $z = \max(F(x) - G(x)) = 0{\cdot}05;\ 0{\cdot}1;\ 0{\cdot}2;\ 0{\cdot}3$. (Table I).

TABLE I

n	k	α	Error of second kind if $z =$			
			0·05	0·1	0·2	0·3
20	6	0·1684		0·6842	0·4332	0·1837
20	7	0·0873		0·8177	0·6163	0·3363
25	7	0·1428		0·7026	0·4154	0·1446
30	8	0·1197		0·7230	0·4039	0·1164
35	9	0·0996		0·7438	0·3961	0·0950
40	9	0·1331		0·6587	0·2640	0·0377
45	10	0·1091	0·8178	0·6890	0·2670	0·0319
50	11	0·0831	0·8430	0·7190	0·2700	0·0265

Before turning to the power function of a pair of statistics, we would like to add a few remarks to our expression $K_n(H_1^{(r)}, \alpha)$ of the power function of the Kolmogorov–Smirnov test.

(a) As we can see from the structure and proof of the formula the power function against the given kind of alternatives can be easily obtained both in the one sample case and two sample case for all test for which the distribution of the test statistic under null hypothesis is known; the idea used is essentially the extension of the method used by Birnbaum [1]. Nevertheless this is simple enough for further considerations and computational work only when the number of linear sections of the distribution function in the alternative hypothesis is small.

(b) If n remains finite and r tends to infinity, then we come essentially to the evaluation of the probabilities of each different array of the sample elements, arrays which belong to the critical region in the case of a given continuous alternative. This kind of expression was given by Hoeffding [3] and considered by Lehmann [5].

3. For r large and for small intervals (z_{j-1}, z_j) we are interested in the mutual order of the sample elements inside the intervals only for very peculiar alternatives. In such cases the two sample problem can be reduced to the null hypothesis

$$H_0' : P\big(\xi \in (z_{j-1}, z_j)\big) = P\big(\eta \in (z_{j-1}, z_j)\big)$$

for $j = 1, 2, \ldots r$. This means that our problem is nothing else but the comparison of two multinomial distributions, which was treated recently by Hoeffding [4].

2. THE POWER FUNCTION OF THE TEST BASED ON A PAIR OF STATISTICS

We shall make use of the notation $nR_{n,n}^+$ for the index of that term in the sequence $(s_0, s_1, s_2, \ldots s_n)$ for which the first maximum occurs, i.e.

$$R_{n,n}^+ = \min_{(i)} \left\{ \frac{i}{n} : s_i = nD_{n,n}^+ \right\}.$$

For deciding between the hypotheses H_0 and $H_1^{(2)}$ (with $g_1 = 0$, for which we have given numerical values of second kind error using the statistic $D_{n,n}^+$), we shall apply now the pair of statistic $(D_{n,n}^+, R_{n,n}^+)$. The author has determined the joint distribution of this pair of statistic under the null hypothesis [8] and with the mentioned alternative [10] as well. These results are now given and then we turn to the numerical values for the second-kind errors:

Introducing the random variable $S_{n,n}^+ = R_{n,n}^+ - D_{n,n}^+$, we make use of the following notations

$$P\left(D_{n,n}^{+} = \frac{k}{n}, S_{n,n}^{+} = \frac{s}{n} \middle| H_0\right) = P_{k,s}^{(n)},$$

and

$$P\left(D_{n,n}^{+} = \frac{k}{n}, S_{n,n}^{+} = \frac{s}{n} \middle| H_1^{(2)}\right) = Q_{k,s}^{(n)}$$

$k = 0, 1, 2, \ldots n$, $s = 1, 2, \ldots n - k$.

Denoting the best critical region of the α-size test, (restricting ourselves to the (k, s)-plane) by $\mathscr{K}_\alpha$ which is defined with the aid of a suitable constant c_α, and which can be written — by virtue of the Neyman–Pearson principle — in the form

$$\mathscr{K}_\alpha = \left\{(k, s) ; \frac{Q_{k,s}^{(n)}}{P_{k,s}^{(n)}} > C_\alpha\right\}$$

with

$$\sum_{(k,s)\in\mathscr{K}_\alpha} P_{k,s}^{(n)} = \alpha .$$

The power function is now

$$K_n(H_1^{(2)}, \alpha) = \sum_{(k,s)\in\mathscr{K}_\alpha} Q_{k,s}^{(n)} .$$

For the probabilities $P_{k,s}^{(n)}$ and $Q_{k,s}^{(n)}$ (see in [10]) we have

$$P_{0,s}^{(n)} = \frac{1}{2(2s-1)(n-s+1)} \frac{\binom{2s}{s}\binom{2n-2s}{n-s}}{\binom{2n}{n}}$$

$s = 1, 2, \ldots, n$,

$$P_{k,s}^{(n)} = \frac{k(k+1)}{(k+2s)(n-s+1)} \frac{\binom{k+2s}{s}\binom{-k+2n-2s}{n-s}}{\binom{2n}{n}}$$

$k = 1, 2, \ldots n$, $s = 0, 1, 2, \ldots n - k$.

Further

$$Q_{0,s}^{(n)} = P_{0,s}^{(n)}(1-z)^n$$

$$s = 1, 2, \ldots n,$$

$$Q_{k,0}^{(n)} = \frac{k+1}{2n-k+1} \frac{\binom{2n-k}{n}}{\binom{2n}{n}} \binom{2n}{k} z^k (1-z)^{n-k}$$

$k = 1, 2, \ldots, n,$

$$Q_{k,s}^{(n)} = \frac{k+1}{n-s+1} \frac{\binom{-k+2n-2s}{n-s}}{\binom{2n}{n}} \sum_{\nu=0}^{k-1} \frac{k-\nu}{n-\nu} \binom{k-\nu+2s}{s} \times$$

$$\times \binom{2n}{n} z^{\nu} (1-z)^{n-\nu}.$$

Using the above formulae for numerical calculation we can compare the second kind errors for $D_{n,n}^{+}$ alone and for $(D_{n,n}^{+}, R_{n,n}^{+})$ respectively. The results are tabulated in Table II.

TABLE II

	Error of first kind		Error of second kind in case of using			
n	$D_{n,n}^{+}$	$(D_{n,n}^{+}, R_{n,n}^{+})$	$z = 0{\cdot}1$ $D_{n,n}^{+}$	$(D_{n,n}^{+}, R_{n,n}^{+})$	$z = 0{\cdot}2$ $D_{n,n}^{+}$	$(D_{n,n}^{+}, R_{n,n}^{+})$
10	0·0739	0·0839	0·8581	0·8071	0·7532	0·6198
	0·2005	0·2038	0·6815	0·5849	0·5215	0·3683
30	0·0675	0·0675	0·8263	0·6149	0·5573	0·2568
		0·0893		0·5636		
	0·1197	0·1197	0·7230	0·5036	0·4039	0·1703
50	0·0562	0·0562	0·8017	0·5105	0·3812	0·1179
	0·0893	0·0893	0·7190	0·4299	0·2700	0·0780
	0·1362	0·1362	0·6133	0·3488	0·1733	0·0473

The computation was carried out on a GIER electronic computer. I am indebted to A. Békéssy and G. Tusnády for their kind help in accomplishing the above calculations.

REFERENCES

1. BIRNBAUM, Z. W., On the Power of a One-sided Test of Fit for Continuous Probability Functions. *Ann. Math. Stat.* **24,** 484–489 (1954).
2. CHAPMAN, D., G., A Comparative Study of Several One-sided Goodness-of-fit Tests. *Ann. Math. Stat.* **29,** 655–674 (1958).
3. HOEFFDING, W., Optimum Non-parametric Tests. *Proc. Sec. Berkeley Symp. Math. Stat. Prob.* Univ. of California Press (1951).
4. HOEFFDING, W., Asymptotically Optimal Tests for Multi-nomial Distributions. *Ann. Math. Stat.* **36,** 369–401 (1965).

5. LEHMAN, E. L., The Power of Rank Tests. *Ann. Math. Stat.* **24,** 23–42 (1953).
6. ROSENBLATT, J., Some Modified Kolmogorov–Smirnov Tests of Approximate Hypothesis and their Properties. *Ann. Math. Stat.* **33,** 513–524 (1962).
7. VINCZE, I., Einige zweidimensionale Verteilungs und Grenzverteilungssätze in der Theorie der geordneten Stichproben. *Publ. Math. Inst. Hung. Acad. Sci.* **2,** 183–209 (1957).
8. VINCZE, I., On Some Joint Distributions and Joint Limiting Distributions in the Theory of Order Statistics. *Publ. Math. Inst. Hung. Acad. Sci.* **4,** 29–47 (1959).
9. VINCZE, I., On Two-sample Tests Based on Order Statistics. *Proc. Fourth Berkeley Symp. Prob. Stat.* **1,** 695–705 (1960).
10. VINCZE, I., On Some Questions Connected with Two-sample Tests of Smirnov-type. *Proc. Fifth Berkeley Symp. Prob. Stat.* (1965).
11. VINCZE, I., A Kolmogorov–Szmirnov és más nemparaméteres próbák erőfüggvényéről. (In Hungarian) *A Magy. Tud. Akad. III (Matematikai és Fizikai) Oszt. Közl.* XV. **2,** 97–105 (1965).

Responsible for publication: Gy. Bernát
Director of the Publishing House of the
Hungarian Academy of Sciences
and of the Academy Press
Responsible editor: E. K. Kállay-Tóth
Technical editor: Á. Garamvölgyi
Jacket designed by: I. MURÁNYI Set in Extended 10/10 pt.
Printed at the Academy Press, Budapest

FURTHER PUBLICATIONS TO BE RECOMMENDED IN THIS LINE

COLLOQUIUM ON APPLICATIONS OF MATHEMATICS TO ECONOMICS, BUDAPEST 1963

Edited by
A. PRÉKOPA

Studies in English, French, German or Russian — 367 pages — 23 figures — 12 tables — 17×24 cm — Cloth

COLLOQUIUM ON THE FOUNDATIONS OF MATHEMATICS, MATHEMATICAL MACHINES AND THEIR APPLICATIONS

Edited by
I. KALMÁR

Studies in English, French, German or Russian — 317 pages 38 figures — 9 tables — 17×24 cm — Cloth

MATHEMATICAL PERIODICALS OF THE HUNGARIAN ACADEMY OF SCIENCES

Acta Mathematica Academiae Scientiarum Hungaricae publishes papers in the field of mathematics in English, German, French or Russian. Published in two issues, making up a volume of some 400 to 500 pages yearly. *Studia Scientiarum Mathematicarum Hungarica* publishes papers on research work pursued in mathematics and its diverse fields of application in science, technology, economics, etc. Papers mainly in English, but also in French, German and Russian. Published in double issues, making up a volume of some 500 pages yearly.

Distributors

KULTURA

Budapest 62. P.O.B. 149